2d and 3

MATH 505a QUALIFYING EXAM Monday, February 9, 2015. One hour and 50 minutes, starting at 5pm.

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1. Let X_n, $n \geq 1$, be independent random variables such that each X_n has Poisson distribution with mean λ_n. Prove that if $\sum_{n\geq 1} \lambda_n = +\infty$, then

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k}{\sum_{k=1}^{n} \lambda_k} = 1$$

in probability.

2. A deck of cards is shuffled thoroughly. Someone goes through all 52 cards, scoring 1 each time 2 cards of the same value are consecutive. For example 9H,8H,7D,6C,7S,7H,7C, scores 2, once due to 7 of spades next to 7 of hearts, and once more 7 of hearts next to 7 of clubs. Write X for the total score.

a) Compute $\mathbb{E}X$.

b) Compute Var$X$.

c) Compute $\mathbb{P}(X = 39)$.

d) In the line below, circle the number that you think is the closest to the value $\mathbb{P}(X = 0)$ and briefly explain your choice.

$$\frac{1}{1000}, \frac{1}{500}, \frac{1}{100}, \frac{1}{50}, \frac{1}{20}, \frac{1}{10}, \frac{1}{5}, \frac{1}{2}.$$

3. Let $S_0, S_1, S_2, \ldots$ be a simple symmetric random walk, i.e. $\mathbb{P}(S_i - S_{i-1} = 1) = \mathbb{P}(S_i - S_{i-1} = -1) = 1/2$, with independent increments. Let $T = \min\{n > 0 : S_n = 0\}$ be the hitting time to zero. Write $\mathbb{P}_a$ for probabilities for the walk starting with $S_0 = a$.

a) What does the reflection principle say about $\mathbb{P}_a(S_n = i, T \leq n)$, for $a > 0$, and $i, n \geq 0$?

b) What does the reflection principle say about $\mathbb{P}_a(S_n \geq i, T > n)$, for $a > 0$, and $i, n \geq 0$? [Hint: telescoping series]

c) For fixed $a > 0$, give asymptotics for $\mathbb{P}_a(T > n)$ as $n \to \infty$. [HINT: Stirling's formula is that $n! \sim \sqrt{2\pi n}\,(n/e)^n$.]

d) Simplify, for fixed $a > 0$,

$$\lim_{n\to\infty} \frac{\mathbb{P}_{a+1}(T > n)}{\mathbb{P}_a(T > n)}.$$

Spring 2015, Applied Probability:

1. For any n, we know that $\sum_{k=1}^{n} X_k \sim \text{Poisson}\left(\sum_{k=1}^{n} \lambda_k\right)$. Let $\varepsilon > 0$ be arbitrary and fixed. Consider

$$P\left(\left|\frac{\sum_{k=1}^{n} X_k}{\sum_{k=1}^{n} \lambda_k} - 1\right| \geq \varepsilon\right) = P\left(\left|\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \lambda_k\right| \geq \varepsilon \sum_{k=1}^{n} \lambda_k\right)$$

$$\leq \frac{E\left[\left|\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \lambda_k\right|^2\right]}{\varepsilon^2 \left(\sum_{k=1}^{n} \lambda_k\right)^2}$$

$$= \frac{\text{Var}\left(\sum_{k=1}^{n} X_k\right)}{\varepsilon^2 \left(\sum_{k=1}^{n} \lambda_k\right)^2}$$

$$= \frac{\left(\sum_{k=1}^{n} \lambda_k\right)}{\varepsilon^2 \left(\sum_{k=1}^{n} \lambda_k\right)^2}$$

$$= \frac{1}{\varepsilon^2 \sum_{k=1}^{n} \lambda_k}$$

Taking limit as $n \to \infty$, since $\sum_{k=1}^{\infty} \lambda_k = \infty$, we obtain

$$\lim_{n \to \infty} P\left(\left|\frac{\sum_{k=1}^{n} X_k}{\sum_{k=1}^{n} \lambda_k} - 1\right| \geq \varepsilon\right) = 0$$

2. Define $X = \sum_{i=1}^{51} 1_{A_i}$ where A_i is the event that i^{th} and $(i+1)^{st}$ are the same

a) $E[X] = \sum_{i=1}^{51} P(A_i)$

$$P(A_i) = \frac{\binom{4}{2}}{\binom{52}{2}} * 13 = \frac{\cancel{4} \cdot 3}{\cancel{2}} \cdot \frac{\cancel{2}}{\cancel{52} * 51} * \cancel{13} = \frac{1}{17}$$

So, $E[X] = \sum_{i=1}^{51} \frac{1}{17} = 3$

b) $\text{Var}(X) = \sum_{i=1}^{51} \text{Var}(1_{A_i}) + 2 \sum_{1 \leq i < j \leq 51} \text{Cov}(1_{A_i}, 1_{A_j})$

$\longrightarrow \text{Var}(1_{A_i}) = P(A_i) - P(A_i)^2 = \frac{1}{17}\left(1 - \frac{1}{17}\right) = \frac{16}{289}$

→ If $j=i+1$: $\quad Cov(1_{A_i}, 1_{A_j}) = P(A_i \cap A_j) - P(A_i)P(A_j)$

$$= \frac{\binom{4}{3}}{\binom{52}{3}} \times 13 - \frac{1}{289}$$

$$= \frac{\cancel{4} \times 3 \times 2}{\cancel{52} \times \cancel{51} \times \cancel{50}} \times \cancel{13} - \frac{1}{289}$$

17 25

$$= \frac{-8}{17^2 * 25}$$

→ If $j \geq i+2$: $\quad Cov(1_{A_i}, 1_{A_j}) = P(A_i \cap A_j) - P(A_i)P(A_j)$

$$= P(A_j|A_i)P(A_i) - P(A_i)P(A_j)$$

$$= \left(\frac{\binom{2}{2}}{\binom{50}{2}} + \frac{\binom{4}{2}}{\binom{50}{2}} * 12 \right) \frac{1}{17} - \frac{1}{289}$$

$$= \left(\frac{\cancel{2}}{\cancel{50} * 49} + \frac{6 * \cancel{2} * 12}{\cancel{50} * 49} \right) * \frac{1}{17} - \frac{1}{289}$$

25 25

$$= \frac{73}{5 * 49} * \frac{1}{17} - \frac{1}{289}$$

$$= \frac{1241}{5 * 49 * (17)^2} - \frac{1}{(17)^2}$$

$$= \frac{996}{5 * 49 * 17^2}$$

So,

$$Var(X) = \frac{\overset{3}{\cancel{51}} \times 16}{\underset{17}{\cancel{289}}} + 2 * 50 * \frac{-8}{17^2 * 25} + 2 * \cancel{49} * \overset{5}{\cancel{25}} * \frac{996}{\cancel{5} * \cancel{49} * (17)^2}$$

$$= \frac{48}{17} - \frac{32}{17^2} + \frac{9960}{17^2}$$

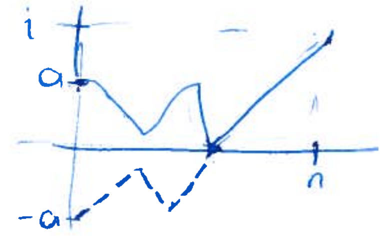
$$= \frac{10744}{289}$$

c) $P(X=39)$ is the probability that all of the numbers $(A, 2, 3, \ldots, J, Q, K)$ are grouped togetger. So,

$$P(X=39) = \frac{13! \, (4!)^{13}}{52!}$$

5243.	.1215927
5244.	.1444267
5245.	.1105418
5246.	.3739063
5247.	.2641677
5248.	.2810553
5249.	.1966646
5250.	.0773446
5251.	.0701672
5252.	.2493338
5253.	.2194934
5254.	.240091
5255.	.[illegible]769193
5256.	.4551708
5257.	.0975352
5258.	.0810651
5259.	.2754133
5260.	.2726697
5261.	.1293996
5262.	.2426514
5263.	.2364552
5264.	.3644556
5265.	.2073167
5266.	.3947792
5267.	.4333436
5268.	.1453843
5269.	.2595214
5270.	.2013097
5271.	.1298397
5272.	.2346963
5273.	.1551946
5274.	.3133776
5275.	.1600424
5276.	.1133135
5277.	.2840638
5278.	.3647679
5279.	.5086656
5280.	.1250206
5281.	.2810553
5282.	.288702
5283.	.1510098
5284.	.2244971
5285.	.3612386
5286.	.2698276
5287.	.278609
5288.	.4752137
5289.	.3030246
5290.	.1182314
5291.	.3513374
5292.	.3284246
5293.	.3244716
5294.	.2455692
5295.	.1133135
5296.	.2738777
5297.	.3495123
5298.	.3996417
5299.	.0791149

$T = \min\{n>0 : S_n = 0\}$



$\mathbb{P}_a(S_n = i,\ T \le n)$ for $a>0$

$= \mathbb{P}_{-a}(S_n = i)$

a'da başlayıp i'de bitenler

— $T_n \le n$ = $-a$'da başlayıp i'de bitenler

$\cup$

— $T_n > n$

$P_a(S_n \ge i,\ T>n)$ $\quad \dfrac{a>0}{i,n\ge 0}$

$\|$

$P_a(S_n \ge i) = 1$

$P_a(S_n \ge i, T>n) = P_a(S_n \ge i) - P_a(S_n \ge i,\ T_n \le n)$

$= P_a(S_n \ge i) - P_{-a}(S_n \ge i)$

$= \sum_{k=i}^{\infty} P_a(S_n = k) - \sum_{k=i}^{\infty} P_a(S_n = k)$

$= \sum_{k=i}^{\infty} [\,P_a($

$= \sum_{k=i}^{\infty} [P_a(S_n = k) - P_a(S_n = k)]$

$\sum_{k=i}^{\infty} \left[\binom{n}{\frac{n+k-a}{2}} \frac{1}{2^n} - \binom{n}{\frac{n+k+a}{2}} \frac{1}{2^n} \right]$

$= \frac{1}{2^n} \sum_{k=i}^{\infty} \left[\binom{n}{\frac{n-a}{2}+\frac{k}{2}} - \binom{n}{\frac{n+a}{2}+\frac{k}{2}} \right]$

$= \frac{1}{2^n} \sum_{k=i}^{n-a}$

West Kolleyu 0219 232 883

Call

call

$n,\ k-a$

$\binom{n}{m}\left(\frac{1}{2}\right)^n$

$m = \frac{n+k-a}{2}$

$\frac{n-a}{2} + \frac{k}{2} \le n$

$n-a+k \le 2n$

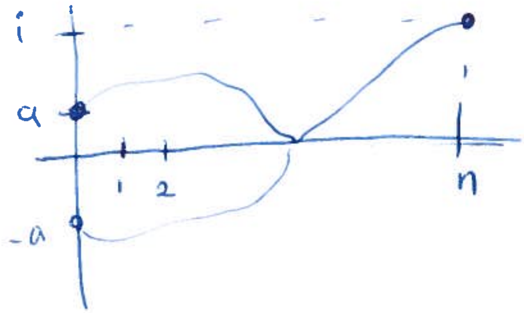
$\boxed{\begin{matrix} k-a \le n \\ k+a \le n \end{matrix}}$

$\frac{n+a}{2} + \frac{k}{2} \le n$

$n+a+k \le 2n$

$k+a \le n$

5300.	.2442482
5301.	.2219872
5302.	.3132033
5303.	.2603506
5304.	.1314668
5305.	.2308419
5306.	.2630866
5307.	.300154
5308.	.3360558
5309.	.0925613
5310.	.4885461
5311.	.3513374
5312.	.2072682
5313.	.215783
5314.	.37079
5315.	.1541944
5316.	.1215927
5317.	.1054445
5318.	.2715307
5319.	.2727768
5320.	.334524
5321.	.1729676
5322.	.1149363
5323.	.2617168
5324.	.3056369
5325.	.2133297
5326.	.1276354
5327.	.2282921
5328.	.2953997
5329.	.2795044
5330.	.2629815
5331.	.3834981
5332.	.2411987
5333.	.1620966
5334.	.318004
5335.	.2853231
5336.	.2097293
5337.	.3147545
5338.	.3541918
5339.	.1506119
5340.	.1062131
5341.	.3071122
5342.	.1370819
5343.	.125888
5344.	.2084723
5345.	.2411987
5346.	.2232402
5347.	.3085905
5348.	.233819
5349.	.2232402
5350.	.28692
5351.	.2826408
5352.	.3090636
5353.	.2060682
5354.	.305967
5355.	.2881854



$$\mathbb{P}_{-a}(S_n = i) =$$

$$\mathbb{P}_a(S_n = i + 2a)$$

$$\sum_{k=i}^{\infty} P_a(S_n = k) - \sum_{k=i}^{\infty} P_a(S_n = k + 2a)$$

$$\sum_{k=i}^{\infty} P_0(S_n = k - a) - \sum_{k=1}^{\infty} P_0(S_n = k + a)$$

$$\sum_{k=i}^{\infty} \left[P_0(S_n = k - a) - P_0(S_n = k + a) \right] = P_0(S_n = i - a) - P_0(S_n = i + a) \qquad k = i$$

$$P_0(S_n = i + 1 - a) - P_0(S_n = i + 1 + a) \qquad k = i + 1$$

$$P_0(S_n = i + 2 - a) - P_0(S_n = i + 2 + a) \qquad k = i + 2$$

$$P_0(S_n = i + 3 - a) - P_0(S_n = i + 3 + a) \qquad k = i + 3$$

$$\vdots$$

$$P_0(S_n = i) \qquad k =$$

$k = i$ $\quad P_a(S_n = i) - P_a(S_n = i + 2a)$

$k = i+1$ $\quad P_a(S_n = i+1) - P_a(S_n = i + 2a + 1)$

$k = i+2$ $\quad P_a(S_n = i+2) - P_a(S_n = i + 2a + 2)$

$k = i + 2a - 1$ $\quad P(S_n = i + 2a - 1) - P(S_n = i + 4a - 1)$

$k = i + 2a$ $\quad P(S_n = i + 2a) - P_a(S_n = i + 4a)$

$$\sum_{k=i}^{i+2a-1} P_a(S_n = k) = P(S_n \geq i, T > n)$$

k

$$\frac{n+a}{2} \leq n \qquad a \leq n$$

$$P_a(T > n) = P_a(S_n > 0, T > n) = P_a(S_n \geq 1, T > n) = \sum_{k=1}^{2a} P_a(S_n = k)$$

$$= \sum_{k=1}^{2a} \binom{n}{\frac{n+k-a}{2}} \frac{1}{2^n} = \sum_{k=1}^{2a} \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi \frac{n+k-a}{2}} \left(\frac{n+k-a}{e}\right)^{\frac{n+k-a}{2}}}$$

MATH 505a QUALIFYING EXAM September 23, 2014. One hour and 50 minutes, starting at 2pm.

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1. Let A and B be two events with $0 < \mathbb{P}(A) < 1$, $0 < \mathbb{P}(B) < 1$. Define the random variables $\xi = \xi(\omega)$ and $\eta = \eta(\omega)$ by

$$\xi(\omega) = \begin{cases} 5 & \text{if } \omega \in A; \\ -7 & \text{if } \omega \notin A; \end{cases} \qquad \eta(\omega) = \begin{cases} 2 & \text{if } \omega \in B; \\ 3 & \text{if } \omega \notin B. \end{cases}$$

True or false: the events A and B are independent if and only if the random variables ξ and η are uncorrelated? If you think this is true, then provide a proof. If you think this is false, then give a counter-example.

2. n people each roll one fair die. For each (unordered) pair of people that get the same number of spots, that number of spots is scored, with S for the total score achieved among the $\binom{n}{2}$ pairs of people. For example, if there are $n = 10$ people, and they roll 1,2,2,2,3,4,4,4,4,6 then $S = 2+2+2+4+4+4+4+4+4$ since there are three pairs of people matching 2 and six $= \binom{4}{2}$ pairs of people scoring 4.

(a) Simplify $\mathbb{E}S$.

(b) Simplify $\mathbb{E}S^2$.

[HINT: Consider S as the sum of $\binom{n}{2}$ random variables $S_{i,j}$, where $S_{i,j}$ is k if persons i and j both roll k, and zero otherwise.]

3. Let X be a standard normal random variable and, for $a > 0$, define the random variable Y_a by

$$Y_a = \begin{cases} X, & \text{if } |X| < a, \\ -X, & \text{if } |X| > a. \end{cases}$$

(a) Verify that Y_a is a standard normal random variable.

(b) Express $\rho(a) = \mathbb{E}(XY_a)$ in terms of the probability density function $\varphi = \varphi(x)$ of X.

(c) Is there a value of a for which $\rho(a) = 0$?

(d) Does the pair (X, Y_a) have a bivariate normal distribution? Explain your reasoning.

Fall 2014, Applied Probability :

1. Suppose A and B are independent.

We can write $\xi = 5\,1_A - 7\,1_{A^c}$ and $\eta = 2\,1_B + 3\,1_{B^c}$. So, we have $\xi\eta = 10\,1_{A\cap B} + 15\,1_{A\cap B^c} - 14\,1_{A^c\cap B} - 21\,1_{A^c\cap B^c}$. Then,

$$\mathbb{E}[\xi\eta] = 10P(A\cap B) + 15P(A\cap B^c) - 14P(A^c\cap B) - 21P(A^c\cap B^c)$$

On the other hand,

$$\mathbb{E}[\xi] * \mathbb{E}[\eta] = (5P(A) - 7P(A^c))(2P(B) + 3P(B^c))$$

$$= 10P(A\cap B) + 15P(A\cap B^c) - 14P(A^c\cap B) - 21P(A^c\cap B^c) \quad \text{by using independence}$$

So, we see that $\mathbb{E}[\xi\eta] = \mathbb{E}[\xi]\mathbb{E}[\eta] < \infty$ which means that, ξ and η are uncorrelated.

Conversely, suppose that ξ and η are uncorrelated. Then define

$$\alpha(\omega) = \xi(\omega) + 7 = \begin{cases} 12 & \text{if } \omega\in A \\ 0 & \text{if } \omega\notin A \end{cases}, \quad \beta(\omega) = \eta(\omega) - 3 = \begin{cases} -1 & \text{if } \omega\in B \\ 0 & \text{if } \omega\notin B \end{cases}$$

Then, we have

$$(\alpha\beta)(\omega) = \begin{cases} -12 & \text{if } \omega\in A\cap B \\ 0 & \text{otherwise} \end{cases}$$

Now, observe that

$$\mathbb{E}[\xi\eta] = \mathbb{E}[(\alpha-7)(\beta+3)] = \mathbb{E}[\alpha\beta] + 3\mathbb{E}[\alpha] - 7\mathbb{E}[\beta] - 21$$

$$\mathbb{E}[\xi]\mathbb{E}[\eta] = (\mathbb{E}[\alpha]-7)(\mathbb{E}[\beta]+3) = \mathbb{E}[\alpha]\mathbb{E}[\beta] + 3\mathbb{E}[\alpha] - 7\mathbb{E}[\beta] - 21$$

Since $\mathbb{E}[\xi\eta] = \mathbb{E}[\xi]\mathbb{E}[\eta]$ $(<\infty)$ we obtain that $\mathbb{E}[\alpha\beta] = \mathbb{E}[\alpha]\mathbb{E}[\beta]$.

We have

$$\mathbb{E}[\alpha\beta] = -12P(A\cap B) \quad \text{and} \quad \mathbb{E}[\alpha]\mathbb{E}[\beta] = 12P(A)(-1)P(B) = -12P(A)P(B)$$

So, we get $P(A\cap B) = P(A)P(B)$ which shows that A and B are independent.

2. Let $S = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} S_{i,j}$ where $S_{i,j}$ is k if persons i and j both roll k, and 0 otherwise.

a) Firstly consider

$$\mathbb{E}[S_{i,j}] = \sum_{m=1}^{6} m\frac{1}{36} = \frac{7}{12}$$

Then, $\mathbb{E}[S] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathbb{E}[S_{ij}] = \frac{7}{12} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 1 = \frac{7}{12} \cdot \frac{n(n-1)}{2} = \frac{7n(n-1)}{24}$

b) Let $R_i := \sum_{j=i+1}^{n} S_{i,j}$ Then $S = \sum_{i=1}^{n-1} R_i$ and

$$\mathbb{E}[S^2] = \sum_{i=1}^{n-1} \mathbb{E}[R_i^2] + 2 \sum_{1 \le i < j \le n-1} \mathbb{E}[R_i R_j]$$

$$\longrightarrow \mathbb{E}[R_i^2] = \sum_{j=i+1}^{n} \mathbb{E}[S_{i,j}^2] + 2 \sum_{i+1 \le j < k \le n} \mathbb{E}[S_{i,j} S_{i,k}]$$

$$* \ \mathbb{E}[S_{i,j}^2] = \sum_{m=1}^{6} m^2 \frac{1}{36} = \frac{91}{36}$$

$$* \ \mathbb{E}[S_{i,j} S_{i,k}] = \sum_{m=1}^{6} m^2 \frac{1}{216} = \frac{91}{216}$$

So, $\mathbb{E}[R_i^2] = (n-i) \frac{91}{36} + (n-i)(n-i-1) \cdot \frac{91}{216} = (n-i) \frac{91}{36} \left(1 + (n-i-1) \cdot \frac{1}{6}\right)$

$\longrightarrow \mathbb{E}[R_i R_j]$ is too complicated!

3. a) $M_{Y_a}(t) = \int_{-\infty}^{\infty} e^{xt} f_{Y_a}(x)\,dx = \int_{-\infty}^{-a} e^{xt} f_{-X}(x)dx + \int_{-a}^{a} e^{xt} f_X(x)dx + \int_{a}^{\infty} e^{xt} f_{-X}(x)dx$

$= \int_{-\infty}^{\infty} e^{xt} f_X(x)\,dx$, since $-X \sim N(0,1)$

$= M_X(t)$ where $X \sim N(0,1)$.

Thus Y_a is a standard normal r.v.

b) Consider

$$XY_a = \begin{cases} X^2, & \text{if } |X| < a \\ -X^2, & \text{if } |X| > a \end{cases}$$

$$\mathbb{E}[XY_a] = \underbrace{\int_{-\infty}^{-a} (-x^2)\varphi(x)\,dx}_{= \int_a^{\infty} -x^2 \varphi(x)\,dx} + \underbrace{\int_{-a}^{a} x^2 \varphi(x)\,dx}_{1 - 2\int_0^{\infty} x^2 \varphi(x)\,dx} + \int_a^{\infty} (-x^2)\varphi(x)\,dx = 1 - 4\int_a^{\infty} x^2 \varphi(x)\,dx$$

c) Consider that $\lim_{a \to \infty} \varphi(a) = 1$ and $\lim_{a \to -\infty} \varphi(a) = -3$. By continuity of $\varphi(a)$, we conclude that $\exists a \in \mathbb{R}$ s.t. $\varphi(a) = 0$.

c) Consider the region $(0,a) \times (-a,0) \in \mathbb{R}^2$, and

$$P((X, Y_a) \in (0,a) \times (-a,0)) = P(X \in (0,a),\ Y_a \in (-a,0))$$

$$= P(Y_a \in (-a,0) \mid X \in (0,a))\, P(X \in (0,a))$$

$$= P(X \in (-a,0) \mid X \in (0,a))\, P(X \in (0,a))$$

$$= 0$$

So, $P((X, Y_a) \in (0,a) \times (-a,0)) = 0$ which would be impossible if (X, Y_a) had bivariate normal distribution.

MATH 505a QUALIFYING EXAM February 6, 2014. One hour and 50 minutes, starting at 5pm.

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1. A person plays a sequence of m games. He wins the nth game with probability a_n independently of the other games. Every time that he wins *two* consecutive games, he is rewarded with \$1; let R be the total reward, so that $0 \le R \le m-1$.

(i) As a function of $a_1, a_2, \ldots, a_m$, give exact expressions for $\mathbb{E}\, R$ and $\mathrm{Var\,R}$.

(ii) Now suppose that $m = 100$ and $a_n = .1$ for all n. Simplify numerically $\mathbb{E}\, R$ and $\mathrm{Var\,R}$.

2. Assume that $X_1, X_2, \cdots$ are independent and identically distributed, each with density

$$f(x) = x/2 \text{ for } 0 < x < 2.$$

For each of the following random variables, simplify the density *or* cumulative distribution function; you may choose either one, for each random variable.

a) $S = X_1 + X_2$.

b) $L = \min(X_1, \ldots, X_{100})$.

c) $R = X_1/X_2$.

d) $M =$ the 10^{th} smallest of $X_1, \ldots, X_{100}$.

3. Let $X_1, X_2, \cdots$ be uncorrelated random variables with $E[X_i] = \mu$ and $var(X_i) \le C < \infty$. If $S_n = X_1 + \cdots + X_n$, show that as $n \to \infty$, $S_n/n \to \mu$ in probability. That is, prove that for any $\varepsilon > 0$,

$$\lim_n \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \ge \varepsilon\right) = 0.$$

Even if you are not comfortable with limits, simply give the best upper bound that you can, of the form

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \ge \varepsilon\right) \le \_\_\_\_\_.$$

Spring 2014, Applied Probability:

1. Define $R = \sum_{i=1}^{m-2} 1_{A_i}$ where A_i is the event that the player wins i^{th} and $(i-1)^{st}$ games.

i) $\mathbb{E}[R] = \sum_{i=1}^{m-1} \mathbb{E}[1_{A_i}] = \sum_{i=1}^{m-1} P(A_i) = \sum_{i=1}^{m-1} a_i a_{i+1}$.

$$var(R) = \sum_{i=1}^{m-1} var(1_{A_i}) + 2 \sum_{1 \le i < j \le m}^{m-1} cov(1_{A_i}, 1_{A_j})$$

$\longrightarrow var(1_{A_i}) = \mathbb{E}[1_{A_i}^2] - (\mathbb{E}[1_{A_i}])^2 = P(A_i) - P(A_i)^2 = a_i a_{i+1}(1 - a_i a_{i+1})$.

$\longrightarrow$ $\underline{j = i+1}$: $cov(1_{A_i}, 1_{A_j}) = \mathbb{E}[1_{A_i} 1_{A_j}] - \mathbb{E}[1_{A_i}]\mathbb{E}[1_{A_j}]$

$$= a_i a_{i+1} a_{i+2} - a_i a_{i+1}^2 a_{i+2}$$

$$= a_i a_{i+1} a_{i+2}(1 - a_{i+1})$$

$\longrightarrow$ $\underline{j \ge i+2}$: $cov(1_{A_i}, 1_{A_j}) = 0$ because of independence.

So,

$$Var(R) = \sum_{i=1}^{m-1} a_i a_{i+1}(1 - a_i a_{i+1}) + 2 \sum_{i=1}^{m-2} a_i a_{i+1} a_{i+2}(1 - a_{i+1})$$

ii) $m = 100$, $a_n = .1$ $\forall n$. Then,

$$\mathbb{E}[R] = \sum_{i=1}^{99} 0.01 = 0.99$$

$$Var(R) = \sum_{i=1}^{99} 0.01 * 0.99 + 2 \sum_{i=1}^{98} 0.001 * 0.9$$

$$= \frac{(99)^2}{10000} + \frac{98 * 18}{10000}$$

$$= 1.1565$$

2. $X_1, \dots, X_n$ are iid with pdf $f(x) = \frac{x}{2}$ for $0 < x < 2$.

a) $F_S(s) = P(S \le s) = P(X_1 + X_2 \le s)$

For $\underline{0 < s \le 2}$:

$$P(X_1 + X_2 \le s) = \int_0^s \int_0^{s-x_2} \frac{x_1 x_2}{4} dx_1 dx_2 = \frac{1}{4} \int_0^s x_2 \int_0^{s-x_2} x_1 dx_1 dx_2 = \frac{1}{4} \int_0^s \frac{x_2 (s-x_2)^2}{2} dx_2$$

$$= \frac{1}{8}\int_0^s (x_2^3 - 2sx_2^2 + s^2x_2)\,dx_2 = \frac{1}{8}\left(\frac{s^4}{4} - \frac{2s^4}{3} + \frac{s^4}{2}\right) = \frac{s^4}{96}$$

For $\underline{2 < s \leq 4}$:

$$P(X_1+X_2 \leq s) = \int_0^2\int_0^{s-2} \frac{x_1x_2}{4}\,dx_1dx_2 + \int_{s-2}^2\int_0^{s-x_1} \frac{x_1x_2}{4}\,dx_2dx_1$$

$$= \frac{1}{4}\int_0^2 x_2 \frac{(s-2)^2}{2}\,dx_2 + \frac{1}{4}\int_{s-2}^2 x_1 \frac{(s-x_1)^2}{2}\,dx_1$$

$$= \frac{(s-2)^2}{4} + \frac{1}{8}\int_{s-2}^2 (x_1^3 - 2sx_1^2 + s^2x_1)\,dx_1$$

$$= \frac{(s-2)^2}{4} + \frac{1}{8}\left[\frac{x_1^4}{4} - \frac{2sx_1^3}{3} + \frac{s^2x_1^2}{2}\right]_{s-2}^2$$

$$= \frac{(s-2)^2}{4} + \frac{1}{8}\left[4 - \frac{16s}{3} + 2s^2 - \frac{(s-2)^4}{4} + \frac{2s(s-2)^3}{3} - \frac{s^2(s-2)^2}{2}\right]$$

$$= \frac{(s-2)^2}{4} + \frac{1}{2} - \frac{2s}{3} + \frac{s^2}{4} - \frac{(s-2)^4}{32} + \frac{s(s-2)^3}{12} - \frac{s^2(s-2)^2}{16}$$

$$= \frac{(s-2)^2}{4}\left[1 - \frac{1}{8} + \frac{s(s-2)}{3} - \frac{s^2}{4}\right] + \frac{s^2}{4} - \frac{2s}{3} + \frac{1}{2}$$

$$= \frac{(s-2)^2}{4}\left[\frac{s^2}{12} - \frac{2s}{3} + \frac{7}{8}\right] + \frac{s^2}{4} - \frac{2s}{3} + \frac{1}{2}$$

So,

$$F_S(s) = \begin{cases} 0 & \text{if } s \leq 0 \\ \frac{s^2}{96} & \text{if } 0 < s \leq 2 \\ \frac{(s-2)^2}{4}\left(\frac{s^2}{12} - \frac{2s}{3} + \frac{7}{8}\right) + \frac{s^2}{4} - \frac{2s}{3} + \frac{1}{2} & \text{if } 2 \leq s < 4 \\ 1 & \text{if } s \geq 4 \end{cases}$$

(probably a mistake in integration provide another solution at the end)

b) $F_L(\ell) = P(L \leq \ell)$

$$= P(\min(X_1, \dots, X_{100}) \leq \ell)$$

$$= 1 - P(\min(X_1, \dots, X_{100}) > \ell)$$

$$= 1 - P(X_1 > \ell, \dots, X_{100} > \ell)$$

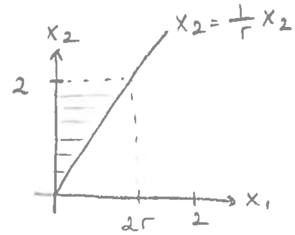
$$= 1 - \prod_{i=1}^n P(X_i > \ell)$$

$$= 1 - \prod_{i=1}^n \int_\ell^2 \frac{x}{2}\,dx$$

$$= 1 - \left(1 - \frac{\ell^2}{4}\right)^n, \quad 0 < \ell < 2.$$

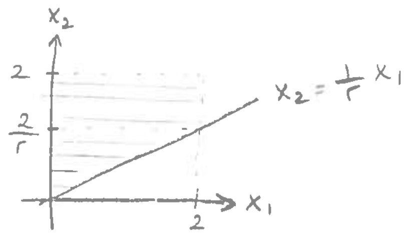
c) $F_R(r) = P\left(\frac{X_1}{X_2} \le r\right) = P(X_1 \le r X_2)$, $0 < r < \infty$

→ If $0 < r < 1$



$$F_R(r) = \int_0^2 \int_0^{rX_2} \frac{x_1 x_2}{4}\, dx_1 dx_2 = \frac{1}{4}\int_0^2 x_2 \frac{r^2 x_2^2}{2}\, dx_2 = \frac{r^2}{8} \frac{x_2^4}{4}\Big|_0^2 = \frac{r^2}{2}$$

→ If $r \geq 1$



$$F_R(r) = \int_0^2 \int_{\frac{x_1}{r}}^2 \frac{x_1 x_2}{4}\, dx_2\, dx_1 = \frac{1}{4}\int_0^2 x_1\left(2 - \frac{x_1^2}{2r^2}\right) dx_1 = \frac{1}{4}\left[x_1^2 - \frac{x_1^4}{8r^2}\right]_0^2$$

$$= \frac{1}{4}\left(4 - \frac{2}{r^2}\right) = 1 - \frac{1}{2r^2}$$

Thus,

$$F_R(r) = \begin{cases} 0 & \text{if } r \leq 0 \\ \frac{r^2}{2} & \text{if } 0 < r < 1 \\ 1 - \frac{1}{2r^2} & \text{if } 1 \leq r < \infty \end{cases}$$

d) We know that

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!\,(n-j)!} \left[F_X(x)\right]^{j-1} f_X(x) \left[1 - F_X(x)\right]^{n-j}$$

So, consider for $0 < x < 2$

$$F_X(x) = \int_0^x \frac{t}{2}\, dt = \frac{x^2}{4}$$

Thus,

$$f_{X_{(10)}}(x) = \frac{100!}{9!\ 90!} \left(\frac{x^2}{4}\right)^9 \frac{x}{2} \left(1 - \frac{x^2}{4}\right)^{90}$$

$$= \frac{100!}{9!\ 90!} \left(\frac{x}{2}\right)^{19} \left(1 - \frac{x^2}{4}\right)^{90}$$

* Consider finding pdf of X_1+X_2 for part (a) by using convolution:

$$f_S(s) = \int_{-\infty}^{\infty} f_{X_1}(x_1)\, f_{X_2}(s-x_1)\, dx_1 = \int_0^2 f_{X_1}(x_1)\, f_{X_2}(s-x_1)\, dx_1 \quad , \quad 0<s<4$$

→ If $0<s<2$.
$$f_S(s) = \int_0^s \frac{x_1}{2}\cdot\frac{s-x_1}{2}\, dx_1 = \frac{1}{4}\left[\frac{s x_1^2}{2} - \frac{x_1^3}{3}\right]_0^s = \frac{s^3}{24}$$

→ If $2\le s<4$.
$$f_S(s) = \int_{s-2}^{s} \frac{x_1}{2}\,\frac{s-x_1}{2}\, dx_1 = \frac{1}{4}\left(2s - \frac{8}{3} - \frac{s(s-2)^2}{2} - \frac{(s-2)^3}{3}\right)$$

So,

$$f_S(s) = \begin{cases} \frac{s^3}{24} & \text{if } 0<s<2 \\ \frac{1}{4}\left(2s-\frac{8}{3}-\frac{s(s-2)^2}{2}-\frac{(s-2)^3}{8}\right) & \text{if } 2\le s<4 \\ 0 & \text{otherwise} \end{cases}$$

3. Consider that $\mathbb{E}\left[\frac{S_n}{n}\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mu$. Now, by using Chebychev's Inequality, we have

$$P\left(\left|\frac{S_n}{n}-\mu\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\,\mathbb{E}\left[\left|\frac{S_n}{n}-\mu\right|^2\right] = \frac{1}{\varepsilon^2}\,\mathrm{var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2\varepsilon^2}\,\mathrm{var}(S_n)$$

Now, observe that

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n}\mathrm{Var}(X_i) + 2\underbrace{\sum_{1\le i<j\le n}\mathrm{Cov}(X_i,X_j)}_{=0\ \forall i\neq j} = \sum_{i=1}^{n}\mathrm{Var}(X_i) \leq Cn$$

So, we get

$$P\left(\left|\frac{S_n}{n}-\mu\right| \geq \varepsilon\right) \leq \frac{1}{n^2\varepsilon^2}\,Cn = \frac{C}{n\varepsilon^2}$$

which is the best upper bound I obtain for the probability. Taking limit as $n\to\infty$, we get

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n}-\mu\right| \geq \varepsilon\right) = 0.$$

3d

MATH 505a QUALIFYING EXAM September , 2013. One hour and 50 minutes

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1.) In an infinite sequence of independent trials, events A, B are mutually exclusive, with $a = \mathbb{P}(A) > 0$ and $b = \mathbb{P}(B) > 0$.
a.) What is the probability that A will occur before B?
b.) In repeated independent tosses of a pair of fair dice, what is the probability that the sum 3 will occur before the sum 7?

2.) Let X and Y be independent, standard normal. Let $W = X + Y$ and $Z = X - Y$.
a.) Show that W and Z are independent.
b.) Simplify $\mathbb{E}(X + 2Y|Z)$.
c.) Simplify $\mathbb{E}(X|X > 0)$.

3.) n balls are placed into d boxes at random, with all d^n possibilities equally likely. Assume $d > 8$. Let X be the number of empty boxes.

a) Calculate and simplify: $\mathbb{E}\,X =$ _____

b) Calculate and simplify: $\mathrm{Var}\,\mathrm{X} =$ _____

c) Let A be the event that boxes 1,2,3,4 are all empty, B be the event that boxes 3,4,5,6 are all empty, and C be the event that boxes 5,6,7,8 are all empty. Compute exactly, $\mathbb{P}(A \cup B \cup C) =$ _____

d) Let D be the event that no box receives more than 1 ball. Fix $a \in (0, 1)$. If both $n, d \to \infty$ together, what relation must they satisfy in order to have $\mathbb{P}(D) \to a$?

Fall 2013, Applied Probability

1. a) Firstly, consider the probability of none of A and B occurs in a trial:

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - a - b.$$

Now, let A_k be the event that A occurs before B for the first time at k^{th} trial. So, $P(A_k) = (1-a-b)^{k-1} a$ and $A = \bigcup_{k=1}^{\infty} A_k$ where A_k's are mutually exclusive. So,

$$P(A) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{\infty} (1-a-b)^{k-1} a = a \cdot \frac{1}{a+b} = \frac{a}{a+b}$$

b) Let A: the event that sum is 3 and B: the event that sum is 7. So,

$P(A) = \frac{1}{18}$ and $P(B) = \frac{1}{6}$. So, the probability that sum 3 will occur before sum 7 is

$$\frac{1/18}{1/18 + 1/6} = \frac{1}{4}$$

2. a) We see that W and Z are linear combinations of independent standard normal r.v. X and Y. So, they are bivariate normal. In this case, to show independence, it is enough to show W and Z are uncorrelated.

$$\left.\begin{array}{l} E[WZ] = E[X^2] - E[Y^2] = 0 \\ E[W]E[Z] = 0 \end{array}\right\} \quad Cov(W,Z) = E[WZ] - E[W]E[Z] = 0$$

Thus, W and Z are independent.

b) Consider that $X + 2Y = X + Y + Y = W + Y$ and since $Y = \frac{W-Z}{2}$, we have

$X + 2Y = W + \frac{W-Z}{2} = \frac{3}{2}W - \frac{1}{2}Z$. So,

$$E[X+2Y \mid Z] = \frac{3}{2}E[W \mid Z] - \frac{1}{2}E[Z \mid Z] = \frac{3}{2}E[W] - \frac{1}{2}Z$$ (by independence and property of conditional expectation)

$$= -\frac{1}{2}Z$$

c) Let A denote the event "$X>0$". Then,

$$E[X \mid X>0] = \frac{E[X 1_A]}{P(A)} = 2E[X 1_A] = 2\int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$u = -\frac{x^2}{2}$

$du = -x\,dx$

$$= -2\int_0^{-\infty} \frac{1}{\sqrt{2\pi}} e^u du = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{0} e^u du = \sqrt{\frac{2}{\pi}}$$

3 Let $X = \sum_{i=1}^{d} 1_{A_i}$ where A_i is the event that ith box is empty.

a) $\mathbb{E}[X] = \sum_{i=1}^{d} P(A_i) = \sum_{i=1}^{d} \left(\frac{d-1}{d}\right)^n = \frac{(d-1)^n}{d^{n-1}}$

b) $\text{Var}\, X = \sum_{i=1}^{d} \text{Var}(1_{A_i}) + 2 \sum_{1 \le i < j \le d} \text{Cov}(1_{A_i}, 1_{A_j})$

$\rightarrow \text{Cov}(1_{A_i}, 1_{A_j}) = \mathbb{E}[1_{A_i} 1_{A_j}] - \mathbb{E}[1_{A_i}]\,\mathbb{E}[1_{A_j}]$

$= P(A_i \cap A_j) - P(A_i)P(A_j)$

$= \left(\frac{d-2}{d}\right)^n - \left(\frac{d-1}{d}\right)^{2n}$

So, $\text{Var}\, X = \sum_{i=1}^{d} \left[\left(\frac{d-1}{d}\right)^n - \left(\frac{d-1}{d}\right)^{2n}\right] + 2 \sum_{1 \le i < j \le d} \left[\left(\frac{d-2}{d}\right)^n - \left(\frac{d-1}{d}\right)^{2n}\right]$

$= \frac{(d-1)^n}{d^{n-1}} - \frac{(d-1)^{2n}}{d^{2n-1}} + \frac{(d-1)(d-2)^n}{d^{n-1}} - \frac{(d-1)^{2n}}{d^{2n-1}}$

c) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

$P(A) = P(B) = P(C) = \left(\frac{d-4}{d}\right)^n$, $P(A \cap B) = P(B \cap C) = \left(\frac{d-6}{d}\right)^n$, $P(A \cap C) = \left(\frac{d-8}{d}\right)^n$

$P(A \cap B \cap C) = \left(\frac{d-8}{d}\right)^n$

So, $P(A \cup B \cup C) = 3\left(\frac{d-4}{d}\right)^n - 2\left(\frac{d-6}{d}\right)^n = \frac{3(d-4)^n - 2(d-6)^n}{d^n}$

d) $P(D) = \frac{1 \cdot (d-1)(d-2)\cdots(d-n+1)}{d^n} = \frac{d(d-1)(d-2)\cdots(d-n+1)(d-n)!}{d^{n+1}(d-n)!} = \frac{d!}{d^{n+1}(d-n)!}$

Using Stirling's formula,

$P(D) \approx \sqrt{2\pi d}\left(\frac{d}{e}\right)^d \frac{1}{d^{n+1}} \frac{1}{\sqrt{2\pi(d-n)}} \frac{1}{\left(\frac{d-n}{e}\right)^{d-n}}$

$= \sqrt{\frac{d}{d-n}} \left(\frac{d}{e}\right)^{d-n} \left(\frac{d}{e}\right)^n \frac{1}{d^{n+1}} \left(\frac{e}{d-n}\right)^{d-n}$

$= \sqrt{\frac{d}{d-n}} \frac{1}{d} e^{-n} \left(\frac{d}{e} \cdot \frac{e}{d-n}\right)^{d-n}$

$= \frac{e^{-n}}{\sqrt{d(d-n)}} \left(\frac{d}{d-n}\right)^{d-n}$

1b 3a, c, e

Math 505a 2013 Spring Qualifying Exam

1. a) Let X and Y be square integrable random variables such that

$$E(X|Y) = Y \quad \text{and} \quad E(Y|X) = X. \tag{1}$$

Show that

$$P(X = Y) = 1. \tag{2}$$

b) Prove that (1) implies (2) under the weakened assumption that X and Y are integrable.

2. Suppose k balls are tossed into n boxes, with all n^k possibilities equally likely. Let D be the number of boxes that contain exactly 2 balls.

a) Compute $p := P($ exactly 2 balls land in box 1$)$.

b) In terms of p, give an exact expression for the mean ED.

c) Compute $r := P($ exactly 2 balls land in box 1 and exactly 2 balls land in box 2$)$.

d) Give an exact expression for the second moment ED^2 in terms of p and r.

e) Compute the variance of D.

3. a) Suppose $g(u) := Eu^S$ is the probability generating function of a non-negative integer valued random variable S satisfying $P(S > 0) > 0$. Let T be distributed as S, conditional on the event $S > 0$. Express $h(u) := Eu^T$, the probability generating function of T, in terms of $g(u)$.

T integer valued, nonnegative?

In parts b) and c) below, N is a nonnegative integer valued random variable with probability generating function $f(u) := Eu^N$, and S is the number of heads in N tosses of a $p \in (0,1)$ coin, with all coin tosses having probability p of coming up heads, independently of each other and of N.

$T | S>0 \overset{d}{=} S$

$$E[u^T] = E[u^T \mid S>0] P\{S>0\} + E[u^T \mid S=0]\, P\{S=0\}$$

$$= g(u) P\{S>0\} + \frac{E[u^T 1_{\{S=0\}}]}{P\{S=0\}} P\{S=0\}$$

$$= g(u) P\{S>0\} + E[u^T 1_{\{S=0\}}]$$

b) Write the probability generating function $g(u) := Eu^S$ of S in a simple form.

c) Now combine parts a) and b): what is the probability generating function h of the number T of heads, in N tosses of a p-coin, conditional on getting at least one head, when N has probability generating function f? Parts d,e) can be worked on even if you are stumped by a,b,c).

$h(u) = \mathbb{E}[u^T]$ is this the same as $\mathbb{E}[u^T | T>0]$ in this question

d) Suppose someone claims that for $\alpha \in (0,1)$, the function

$$f(u) := 1 - (1-u)^{\alpha}$$

is a probability generating function of a nonnegative, non constant integer valued random variable N. What properties of f must you check? Is the hypothesis $\alpha > 0$ used? What happens in the cases $\alpha = 0, \alpha = 1$ and $\alpha > 1$?

e) Combine parts a)-d), that is suppose $\alpha \in (0,1)$, N has the generating function $f(u) := 1 - (1-u)^{\alpha}$, and T is the number of heads in N tosses of a p-coin, conditional on getting at least one head. Do N and T have the same distribution?

$P(N=k) = \dfrac{f^{(k)}(0)}{k!} \Rightarrow P(N=0) = 0$

$P(N=1) = \alpha$

$P(N=k) = \dfrac{1}{k!} \prod_{i=0}^{k-1} (\alpha - i), \quad k \geq 1$

- "$\alpha > 0$" is requiered since $P(N=1) = \alpha$
- If $\alpha = 0$ $f(u)$ is not a prob. g. func.
- If $\alpha = 1$ $P(N=1) = 1$
- If $\alpha > 1$ $f(u)$ is not a prob. g func since $P(N=1) = \alpha$

Spring 2013, Applied Probability

1. a) $\hookrightarrow$ $\underline{Var(X|Y) = 0}$:

$$Var(X) = \mathbb{E}[Var(X|Y)] + Var(\mathbb{E}[X|Y])$$
$$= \mathbb{E}[Var(X|Y)] + Var(Y)$$
$$= \mathbb{E}[Var(X|Y)] + \mathbb{E}[Var(Y|X)] + Var(\mathbb{E}[Y|X])$$
$$= \mathbb{E}[Var(X|Y) + Var(Y|X)] + Var(X)$$

Since $Var(X) < \infty$ (because of assumption) we get $\mathbb{E}[\underbrace{Var(X|Y) + Var(Y|X)}_{\geq 0}] = 0$

$\Rightarrow Var(X|Y) + Var(Y|X) \geq 0 \Rightarrow Var(X|Y) = 0$ and $Var(Y|X) = 0$.

$\longrightarrow$ $\underline{Var(X) = Var(Y)}$:

$$0 = Var(X|Y) = \mathbb{E}[X^2|Y] - (\mathbb{E}[X|Y])^2 = \mathbb{E}[X^2|Y] - Y^2$$

So, we have $\mathbb{E}[X^2|Y] = Y^2$ and taking expectation of both sides, we obtain $E[X^2] = E[Y^2]$. This also gives us $Var(X) = Var(Y)$ since $\mathbb{E}[X] = \mathbb{E}[Y]$ by assumption $(\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[Y])$.

$\longrightarrow$ $\underline{\mathbb{E}[XY] = \mathbb{E}[X^2]}$:

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY|X]] = \mathbb{E}[X\,\mathbb{E}[Y|X]] = \mathbb{E}[X^2]$$

$\longrightarrow$ $\underline{P(X=Y) = 1}$:

$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$
$$= 2Var(X) - 2\mathbb{E}[XY] + 2\mathbb{E}[X]\,\mathbb{E}[Y]$$
$$= 2\mathbb{E}[X^2] - 2(\mathbb{E}[X])^2 - 2\mathbb{E}[X^2] + 2(\mathbb{E}[X])^2$$
$$= 0$$

So, $P(X-Y=c) = 1$ i.e. $X = Y + c$ a.e. Since $\mathbb{E}[X] = \mathbb{E}[Y]$, we must have $c=0$.
Thus, $P(X=Y) = 1$

2. $D = \#$ of boxes that contain exactly 2 balls

a) $X = \#$ of balls in box 1, and $P(\text{any ball lands in box } 1) = \frac{1}{n}$. So,
$X \sim \text{binomial}(k, \frac{1}{n})$ and $p = P(X=2) = \binom{k}{2}\left(\frac{1}{n}\right)^2\left(\frac{n-1}{n}\right)^{k-2} = \binom{k}{2} \cdot \frac{(n-1)^{k-2}}{n^k}$.

b) $D = \sum_{i=1}^{n} 1_{D_i}$ where D_i is the event that i^{th} box contains exactly 2 balls. Then, $\mathbb{E}[D] = \sum_{i=1}^{n} P(D_i)$ where $P(D_i) = \binom{k}{2} \frac{(n-1)^{k-2}}{n^k}$ by part (a).

So,

$$\mathbb{E}[D] = \sum_{i=1}^{n} \binom{k}{2} \frac{(n-1)^{k-2}}{n^k} = \sum_{i=1}^{n} p = np$$

c) Firstly, let $X =$ # of balls land in first box. So, $X \sim \text{Binomial}(k, \frac{1}{n})$, then letting $Y =$ # of balls land in second box, we see that $Y|X \sim \text{Binomial}(k-X, \frac{1}{n-1})$. So,

$$P(X=2, Y=2) = P(Y=2 \mid X=2)\, P(X=2)$$

$$= \binom{k-2}{2} \frac{1}{(n-1)^2} \left(\frac{n-2}{n-1}\right)^{k-4} \binom{k}{2} \frac{1}{n^2} \left(\frac{n-1}{n}\right)^{k-2}$$

$$= \binom{k}{2}\binom{k-2}{2} \frac{(n-2)^{k-4}}{n^k}$$

(*choose 2 then choose 2 × $\frac{\text{all possible in}}{\text{all possible wh}}$

So, $r = \binom{k}{2}\binom{k-2}{2} \frac{(n-2)^{k-4}}{n^k}$

d)
$$\mathbb{E}[D^2] = \sum_{i=1}^{n} \mathbb{E}[1_{D_i}^2] + \sum_{1 \le i < j \le n} \mathbb{E}[1_{D_i} 1_{D_j}]$$

$$= \sum_{i=1}^{n} P(D_i) + \sum_{1 \le i < j \le n} P(D_i \cap D_j)$$

$$= \sum_{i=1}^{n} p + \sum_{1 \le i < j \le n} r$$

$$= np + \frac{(n-1)n}{2} r$$

e) $\text{Var}(D) = \mathbb{E}[D^2] - \mathbb{E}[D]^2 = np + \frac{(n-1)n}{2} r - n^2 p^2 = np(1-np) + \frac{(n-1)n}{2} r$

3. a)
$$\mathbb{E}[u^T] = \mathbb{E}[u^T \mid S>0]\, P\{S>0\} + \mathbb{E}[u^T \mid S=0]\, P\{S=0\}$$

$$= g(u)\, P\{S>0\} + \frac{\mathbb{E}[u^T 1_{\{S=0\}}]}{P\{S=0\}} P\{S=0\}$$

$$= g(u)\, P\{S>0\} + \mathbb{E}[u^T 1_{\{S=0\}}].$$

b) We have $S \sim \text{Binomial}(N,p)$. Then,

$$\mathbb{E}[u^S \mid N=n] = \sum_{k=0}^{n} u^k P(S=k \mid N=n)$$
$$= \sum_{k=0}^{n} u^k \binom{n}{k} p^k (1-p)^{n-k}$$
$$= \sum_{k=0}^{n} \binom{n}{k} (pu)^k (1-p)^{n-k}$$
$$= (1-p+pu)^n$$

So, $\mathbb{E}[u^S \mid N] = (1-p+pu)^N$. Then,

$$\mathbb{E}[u^S] = \mathbb{E}\left[\mathbb{E}[u^S \mid N]\right] = \mathbb{E}\left[(1-p+pu)^N\right] = f(1-p+pu).$$

c) T = # of heads in N tosses of a p-coin, conditional on getting at least one head

$$m(u) = \mathbb{E}[u^T \mid T>0] = \frac{\mathbb{E}[u^T 1_{\{T>0\}}]}{P\{T>0\}} \quad \text{where } T \sim \text{Binomial}(N,p).$$

$$\rightarrow P\{T>0\} = \sum_{n=1}^{\infty} P\{T>0 \mid N=n\} P\{N=n\} = \sum_{n=1}^{\infty} [1-(1-p)^n] P\{N=n\} = \underbrace{\sum_{n=1}^{\infty} P\{N=n\}}_{\text{since } =1<\infty} - \underbrace{\sum_{n=1}^{\infty} (1-p)^n P\{N=n\}}_{<\infty}$$

$$= 1 - \mathbb{E}[(1-p)^N] = 1 - f(1-p).$$

$$\rightarrow \mathbb{E}[u^T 1_{\{T>0\}}] = \sum_{k=0}^{\infty} u^k 1_{\{T>0\}} P\{T=k\} = \sum_{k=1}^{\infty} u^k P\{T=k\} = \sum_{k=0}^{\infty} u^k P\{T=k\} - P\{T=0\}$$

$$= h(u) - P\{T=0\} = h(u) - 1 + P\{T>0\} = h(u) - 1 + 1 - f(1-p)$$
$$= h(u) - f(1-p)$$

Thus, $m(u) = \dfrac{h(u) - f(1-p)}{1 - f(1-p)}$.

d) To have $f(u) = 1-(1-u)^\alpha$ as a probability generating function of N, we must have

i) $f(1)=1$ since $f(u) = \mathbb{E}[u^N] = \sum_{k=0}^{\infty} u^k P\{N=k\}$ implies $f(1)=1$. For the given function $f(u)$, $f(1) = 1-(1-1)^\alpha = 1$, if $\alpha \geq 0$, this condition is satisfied.

Note that if $\alpha=0$, then $f(u)\equiv 1$ which is not a probability generating function. So, let $\alpha>0$

ii) $P\{N=k\} = \dfrac{f^{(k)}(0)}{k!}$, i.e. f must be infinitely many times differentiable and $0 < \dfrac{f^{(k)}(0)}{k!} < 1$ and $\displaystyle\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} = 1$.

The given function f is infinitely many times differentiable. Then, consider

$$f'(u) = \alpha(1-u)^{\alpha-1} \longrightarrow P\{N=1\} = f'(0) = \alpha$$

$$f''(u) = -\alpha(\alpha-1)(1-u)^{\alpha-2} \longrightarrow P\{N=2\} = \frac{\alpha(1-\alpha)}{2}$$

$$f'''(u) = \alpha(\alpha-1)(\alpha-2)(1-u)^{\alpha-3} \longrightarrow P\{N=3\} = \frac{\alpha(\alpha-1)(\alpha-3)}{6}$$

$$\vdots$$

$$f^{(k)}(u) = (-1)^{k-1}\prod_{i=1}^{k-1}(\alpha-i)(1-u)^{\alpha-k} \longrightarrow P\{N=k\} = \frac{(-1)^{k-1}}{k!}\prod_{i=1}^{k-1}(\alpha-i)$$

Since $P\{N=1\}$ we must have $\alpha \leq 1$

If $\alpha = 1$, $P\{N=0\} = 0$ and $P\{N=1\} = 1$

So, we must have $0 < \alpha \leq 1$ to have $f(u)$ as a probability generating function

If $\alpha > 1$, $f(u)$ cannot be a probability generating function

Math 505a 2012 Fall Qualifying Exam

1. In the Polya Urn model, $w \geq 1$ white balls and $b \geq 1$ black balls are placed in an urn at time 0, and at times $1, 2, \ldots$ a ball is chosen uniformly from the urn independent of the past, and replaced back into the urn with one additional ball of the same color.

a. A vector $(X_1, \ldots, X_n)$ of random variables is said to be exchangeable

$$(X_1, \ldots, X_n) =_d (X_{\pi(1)}, \ldots, X_{\pi(n)}) \quad \text{for all permutations } \pi$$

where $=_d$ denotes equality in distribution. If X_i is the indicator that a white ball is drawn from the urn at time i, prove that $(X_1, \ldots, X_n)$ is exchangeable.

b. Find the mean and variance of $S_n = X_1 + \cdots + X_n$, the total number of white balls added to the urn up to time n.

2. With a and b positive numbers, a needle of length $l \in (0, \min(a, b)]$ is dropped randomly on a rectangular grid consisting of an infinite number of parallel lines distance a apart, and, perpendicular to these, an infinite number of parallel lines distance b apart. Let A and B, respectively, be the events that the needle intersects the group of lines at distance a and b apart.

a. Show $P(A) = \frac{2l}{a\pi}$ and $P(B) = \frac{2l}{b\pi}$. Hint: The angle θ giving the orientation of the needle might be taken as uniform from $[0, 2\pi)$, but by symmetry, one may assume that the angle is uniformly taken from $[0, \pi/2)$.

b. Determine $P(A \cap B)$ and verify that A and B are strictly negatively correlated, that is, that $P(A \cap B) < P(A)P(B)$.

3. A total of k boys and $n - k$ girls sit around a circular table, with all $\cancel{n!}$ $(n-1)!$ arrangements equally likely. Compute the mean and variance of the number Y of pairs of boy/girl neighbors. Note: In the circular arrangement GBGGBB, since the first G and last B are neighbors, $Y = 4$.

Fall 2012, Applied Probability

1. a) Consider probability $P(X_1 = x_1, \ldots, X_n = x_n)$ where $x_1, \ldots, x_n \in \{0,1\}$.
Then, this probability is the probability of having exactly $\sum_{i=1}^{n} x_i$ white balls among chosen n balls. So, the order of white balls does not affect the probability. Thus for any π, $P(X_{\pi(1)} = x_1, \ldots, X_{\pi(n)} = x_n)$ represents the probability of getting $\sum_{i=1}^{n} x_i$ white balls among chosen n balls.
Thus,

$$(X_1, \ldots, X_n) \stackrel{d}{=} (X_{\pi(1)}, \ldots, X_{\pi(n)})$$

for all permutations π.

b) Let $S_n = X_1 + \cdots + X_n$ represent the total number of balls added to the urn up to time n. Here,

$$X_i = \begin{cases} 1 & \text{if a white ball added at time } i \\ 0 & \text{otherwise} \end{cases}$$

$$\rightarrow \mathbb{E} S_n = \sum_{i=1}^{n} \mathbb{E} X_i = \sum_{i=1}^{n} P\{X_i = 1\}$$

Now, consider

- $P(X_1 = 1) = \dfrac{w}{w+b}$

- $P(X_2 = 1) = P(X_2 = 1 | X_1 = 1) P(X_1 = 1) + P(X_2 = 1 | X_1 = 0) P(X_1 = 0)$ $\quad (\star)$

$$= \frac{w+1}{w+b+1} \cdot \frac{w}{w+b} + \frac{w}{w+b+1} \cdot \frac{b}{w+b}$$

$$= \frac{w(w+b+1)}{(w+b)(w+b+1)}$$

$$= \frac{w}{w+b}$$

- Observe above that we can replace 2 by $i+1$ and 1 by i in $(\star)$ above, since $P(X_1 = 1) = P(X_2 = 1)$ (inductively). Then, we obtain

$$P(X_i = 1) = \frac{w}{w+b}$$

Thus, we obtain

$$\mathbb{E}[S_n] = \frac{nw}{w+b}$$

$$\longrightarrow Var(S_n) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{1 \le i < j \le n} Cov(X_i, X_j)$$

- $Var(X_i) = P\{X_i = 1\} - P\{X_i = 1\}^2 = \frac{w}{w+b}\left(1 - \frac{w}{w+b}\right) = \frac{wb}{(w+b)^2}$

- $Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j] = P\{X_i = 1, X_j = 1\} - P\{X_i = 1\}P\{X_j = 1\}$

  - For $i<j$,

$$\left.\begin{aligned} P\{X_i = 1, X_j = 1\} &= P\{X_j = 1 \mid X_i = 1\} P\{X_i = 1\} \\ &= P\{X_2 = 1 \mid X_1 = 1\} P\{X_1 = 1\} \end{aligned}\right\}$$ by the same reasoning in part (a).

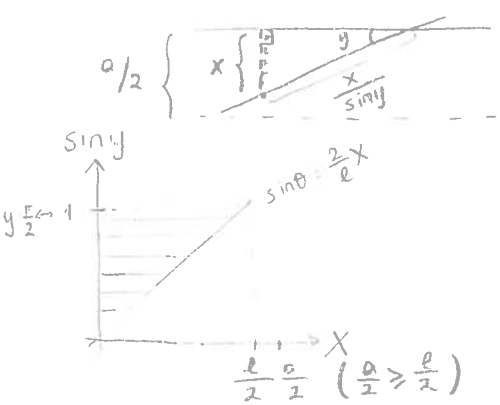
$$= \frac{w+1}{w+b+1} \; \frac{w}{w+b}$$

$$Cov(X_i, X_j) = \frac{(w+1)w}{(w+b+1)(w+b)} - \frac{w^2}{(w+b)^2} = \frac{w}{w+b}\left(\frac{w+1}{w+b+1} - \frac{w}{w+b}\right)$$

$$Var(S_n) = \frac{nwb}{(w+b)^2} + \frac{n(n-1)w}{w+b}\left(\frac{w+1}{w+b+1} - \frac{w}{w+b}\right).$$

2 a) θ : the accute angle between the needle and the horizontal lines (wlog)
X : the distance between the midpoint of the needle and horizontal lines

Assume also wlog that the distance between horizontal lines is a and the distance between vertical lines is b



$\theta \sim U[0, \frac{\pi}{2})$ and $X \sim U[0, \frac{a}{2})$ by symmetry

$$\Rightarrow P(A) = P\left(\frac{X}{\sin y} < \frac{\ell}{2}\right)$$

$$= \int_0^{\pi/2} \int_0^{\ell \sin y/2} f_X(x) f_\theta(y)\, dx\, dy = \frac{4}{a\pi} \int_0^{\pi/2} \frac{\ell}{2} \sin y \, dy$$

$$= \frac{2\ell}{a\pi}\left[-\cos y\right]_0^{\pi/2} = \frac{2\ell}{a\pi}$$

By changing horizontal's to vertical's, a's to b's and keeping θ the same, the above argument works for $P(B)$ in exactly the same way. Thus, we get $P(B) = \frac{2\ell}{b\pi}$.

b) Now, consider

$$P(A \cap B) = P\left(\frac{X}{\cos\theta} < \frac{\ell}{2}, \frac{Y}{\sin\theta} < \frac{\ell}{2}\right)$$

$$= P\left(X < \frac{\ell\cos\theta}{2}, Y < \frac{\ell\sin\theta}{2}\right)$$

$$= \int_0^{\pi/2} \int_0^{\frac{\ell\sin\theta}{2}} \int_0^{\frac{\ell\cos\theta}{2}} \frac{2}{a}\,\frac{2}{b}\,\frac{2}{\pi}\, dx\, dy\, d\theta$$

$$= \frac{8}{ab\pi} \int_0^{\pi/2} \frac{\ell^2 \sin 2\theta}{8}\, d\theta$$

$$= \frac{\ell^2}{ab\pi}\left[-\frac{\cos 2\theta}{2}\right]_0^{\pi/2}$$

$$= \frac{\ell^2}{ab\pi}\left[\frac{1}{2} + \frac{1}{2}\right]$$

$$= \frac{\ell^2}{ab\pi}$$

In this case, $P(A \cap B) < P(A)P(B) \iff \frac{\ell^2}{ab\pi} < \frac{4\ell^2}{ab\pi^2} \iff 1 < \frac{4}{\pi}$ which is always true. So, we have $P(A \cap B) < P(A)P(B)$

$$f_{U_2+U_3}(t) = \int_0^1 f_{U_2}(x)\, f_{U_3}(t-x)\,dx$$

$0<x<1$ $\quad$ $t-1<x<t$ $\quad$ $-t<-x<1-t$

$$f_{U_3}(t-x) = \begin{cases} 1 & \text{if } 0<t-x<1 \\ 0 & \text{otherwise} \end{cases}$$

$x+y=t$

$$= \int_0^1 f_{U_3}(t-x)\,dx$$

$0<y<1$ $\quad$ $0<x+y<2$ $\quad$ $0<t<2$

$0<t-x<1$

$t-1<x<t$

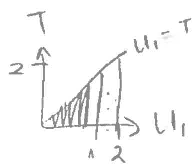
If $\underline{0<t<1}$: $\quad = \int_0^t dx = t$ $\quad$ $0<t<1$

If $\underline{1\le t<2}$: $\quad = \int_{t-1}^1 dx = 2-t$

$$f_{U_2+U_3}(t) = \begin{cases} t, & 0<t<1 \\ 2-t, & 1\le t<2 \\ 0 & \text{otherwise} \end{cases}$$

$$P(U_1 > \underbrace{U_2+U_3}_{T}) = \int_0^1\int_0^x f_{U_1}(x)\, f_T(t)\,dt\,dx = \int_0^1\int_0^x t\,dt\,dx = \frac{1}{2}\int_0^1 x^2\,dx$$

$$= \frac{1}{6}\left[x^3\right]_0^1 = \frac{1}{6}$$



$0<x<1$

$0<t<x \Rightarrow 0<t<x<1$

Number the places $1,2,\dots,n$ $\quad$ Define for $i=1,2,\dots,n-1$

$$1_i = \begin{cases} 1 & \text{if neighbors sitting } i \text{ and } i+1 \text{ are boy/girl (girl/boy)} \\ 0 & \text{otherwise} \end{cases}$$

~~A_i the event that~~

and define 1_n in the same way for the people sitting at n and 1.

$$Y = \sum_{i=1}^n 1_{A_i} \qquad P(A_i) = \frac{2(n-k)k}{n(n-1)} \qquad \boxed{E[Y] = \frac{2(n-k)k}{n-1}}$$

$$\operatorname{Var}(1_{A_i}) = E[1_{A_i}] - (E[1_{A_i}])^2 = \frac{2(n-k)k}{n(n-1)}\left(1 - \frac{2(n-k)k}{n(n-1)}\right)$$

$$\operatorname{Var} Y = \sum_{i=1}^n \operatorname{var}(1_{A_i}) + 2\sum_{1\le i<j\le n} \operatorname{Cov}(1_{A_i}, 1_{A_j})$$

$$\operatorname{Cov}(1_{A_i}, 1_{A_j}) = P(A_i\cap A_j) - P(A_i)P(A_j)$$

$\underline{j=i+1}$: $\quad$ (BGB, GBG)

$$P(A_i\cap A_j) = \frac{(n-k)k(k-1)}{n(n-1)(n-2)} + \frac{(n-k-1)k(n-k)}{n(n-1)(n-2)} = \frac{k(n-k)[\cancel{k}-1+n-\cancel{k}-1]}{n(n-1)(\cancel{n-2})} = \frac{k(n-k)}{n(n-1)}$$

~~k-1+n-k-1~~ $\quad$ n of them

+1 from the next case

$P(BGB) = P(3.B\mid BG)P(BG)$

$= \frac{k-1}{n-2}\cdot P(2.G\mid B)P(B)$

$= \frac{k-1}{n-2}\cdot\frac{n-k}{n-1}\cdot\frac{k}{n}$

$$\boxed{A_1\cap A_2,\ A_2\cap A_3,\ A_3\cap A_4,\ \dots,\ A_{n-1}\cap A_1} \to (n-1) \text{ of them}$$

(n-1, n, 1)

$$\operatorname{Cov}(1_{A_i}, 1_{A_j}) = \frac{k(n-k)}{n(n-1)} - \frac{4(n-k)^2k^2}{n^2(n-1)^2} = \frac{k(n-k)}{n(n-1)}\left(1 - \frac{4k(n-k)}{n(n-1)}\right)$$

$P(GBG) = \frac{n-k-1}{\dots}\ \frac{k}{\dots}\ \frac{n-k}{\dots}$

DEPARTMENT OF MATHEMATICS – SCREENING EXAMINATION APPLICATION

The Fall 2015 Graduate Screening Examinations will be given the week of

SEPTEMBER 14-29, 2015

Please return completed application and availability forms to Amy (amy@usc.edu) by

12 noon, Wednesday, September 2, 2015

***Important reminder: It is the responsibility of the degree candidate to check the USC catalogue and his/her STARS report (accessible via oasis) to assure that all degree requirements are fulfilled.*

****

Your name: ______________________ Your USC ID#: ______________________

USC email: ______________________

Degree objective (e.g., MS, MA, PHD): ______________________

Degree program (e.g., MATH, AMAT, STAT): ______________________

Please list all departmental screening, qualifying, or comprehensive exams that you have taken in the past and the results from those exams:

Select from the following list, (√) all the exams that you wish to take during the Fall 2015 examination period.

____ ALGEBRA (FORMERLY KNOWN AS 510AB)
____ APPLIED PROBABILITY (FORMERLY KNOWN AS 505A)
____ COMPLEX ANALYSIS (FORMERLY KNOWN AS 520)
____ DIFFERENTIAL EQUATIONS (ORDINARY AND PARTIAL: FORMERLY KNOWN AS 555A & 565A)
____ GEOMETRY AND TOPOLOGY (FORMERLY KNOWN AS 535A & 540)
____ NUMERICAL ANALYSIS (FORMERLY KNOWN AS 502A)
____ NUMERICAL ANALYSIS B (FORMERLY KNOWN AS 502B)
____ PROBABILITY (FORMERLY KNOWN AS 507A)
____ REAL ANALYSIS (FORMERLY KNOWN AS 525A)
____ STATISTICS (FORMERLY KNOWN AS 541A)
____ STATISTICS B (FORMERLY KNOWN AS 541B)

$j \geq i+2$: $\quad P(A_i \cap A_j) = P(A_j \mid A_i)\, P(A_i)$

$A_1 \cap A_3,\ A_1 \cap A_4, \dots, A_1 \cap A_{n-1}$ $\quad A_1 \cap A_n$ previous case

$A_2 \cap A_4,\ A_2 \cap A_5, \dots, A_2 \cap A_{n-1}$ $\quad A_2 \cap A_n$

$A_3 \cap A_5,\ A_3 \cap A_6, \dots, A_3 \cap A_{n-1}, A_3 \cap A_n$

$$= \frac{2(n-k-1)(k-1)}{(n-2)(n-3)} \cdot \frac{2(n-k)k}{n(n-1)}$$

$$\mathrm{Cov}(1_{A_i}, A_j) = \frac{4k(k-1)(n-k)(n-k-1)}{n(n-1)(n-2)(n-3)} - \frac{4(n-k)^2k^2}{n^2(n-1)^2}$$

$A_{n-2} \cap A_n$ ✓

$$= \frac{4(n-k)k}{n(n-1)} \left(\frac{(n-k-1)(k-1)}{(n-2)(n-3)} - \frac{(n-k)k}{n(n-1)} \right)$$

$k=2 \longrightarrow n-3 \qquad (n-k-1)$

$k=3 \longrightarrow n-4 \qquad (n-k-1$

$\vdots$

$k=n-2 \qquad 1$

$$\frac{(n-3)(n-2)}{2} + (n-3) \qquad (n-3)\left[\frac{n-2}{2}+1\right] = \frac{n(n-3)}{2}$$

of them

$$\text{So, } \mathrm{Var}(Y) = \frac{2(n-k)k}{(n-1)}\left(1 - \frac{2(n-k)k}{n(n-1)}\right) + 2\left[n * \frac{k(n-k)}{n(n-1)}\left(1 - \frac{4k(n-k)}{n(n-1)}\right) + \frac{n(n-3)}{2} * \frac{4(n-k)k}{n(n-1)}\left(\frac{(n-k-1)(k-1)}{(n-2)(n-3)} - \frac{(n-k)k}{n(n-1)}\right)\right]$$

Name:

Math 114x - Spring 2015 Quiz 2
January 29, 2015

Question 1: (a) Find the average, the deviations from average and the standard deviation of the following list: 9, 8, 12, 10, 6.

average:
deviations from average:
standard deviation:

(b) Which numbers on the list are withing one standard deviations of average?

Question 2: You are looking at a computer printout of 100 test scores, which have been converted to standard units. The first 10 entries are

-6.2, 3.5, 1.2, -0.13, 4.3, -5.1, -7.2, -11.3, 1.8, 6.3

Does the printout look reasonable, or is something wrong with the computer?

MATH 505a QUALIFYING EXAM January 31, 2012

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1.) One hundred and one passengers bought tickets on a 101-seat carriage. One seat was reserved for each passenger. The first 100 passengers took the seats at random so that all 101! possible seating arrangements (with one empty seat) are equally likely. The last passenger insisted on taking the assigned seat. If that seat is occupied, then the passenger in that seat has to move to the corresponding assigned seat, and so on. Compute the expected value of the number M of passengers who have to change their seats. [HINTS: one method is to use a recursion in n, for n in the role of 101, for the expectation, without knowing the distribution of M. Another method is to find the distribution of M explicitly.]

2.) Suppose $\mathbb{P}(X = k) = p_k$ and $p_1 + p_2 + \cdots = 1$. Suppose that $X, X_1, X_2, \ldots, X_n$ are independent and identically distributed. Let $S = \sum_{1 \le i < j \le n} 1(X_i = X_j)$ be the number of matching (unordered) *pairs*, and let $T = \sum_{1 \le i < j < k \le n} 1(X_i = X_j = X_k)$ be the number of matching (unordered) *trios*. For $r = 1, 2, 3, \ldots,$, let $f_r = \sum p_i^r$, so that $f_1 = 1$.
a) Give a simple expression for $\mathbb{E}\, S$ in terms of n, f_2.
b) Give a simple expression for $\mathbb{E}\, T$ in terms of n, f_3.
c) Give a simple expression for Var S in terms of n, f_2, f_3, f_4.

3.) Let X, X_1, X_2, X_3, X_4 be independent standard exponentially distributed random variables, so that $\mathbb{P}(X > x) = e^{-x}$ for $x > 0$. Write $S_n = X_1 + \cdots + X_n$. The goal is to show that the triple $(S_1/S_4, S_2/S_4, S_3/S_4)$ is distributed like the order statistics of three independent standard uniform (0,1) random variables.
a) Give a reason why the density of S_4 is $f(t) = t^3 e^{-t}/6$ for $t > 0$. You may either quote the density for the Gamma family in general, or you may argue about the time of the fourth arrival in a standard, rate 1 Poisson process, or you may carry out the four-fold convolution!
b) With (U_1, U_2, U_3) distributed uniformly in the cube $(0,1)^3$, and $U_{[i]}$ defined to be the ith smallest of U_1, U_2, U_3, show why the density of $(U_{[1]}, U_{[2]}, U_{[3]})$ is $g(x, y, z) = 6$, on the region $0 < x < y < z < 1$.
c) Show, with detail, why the triple $(S_1/S_4, S_2/S_4, S_3/S_4)$ is distributed like the order statistics of three independent standard uniform (0,1) random variables, AND that this triple is independent of S_4. This should include calculation of both a 4 by 4 Jacobian matrix, and calculation of its determinant, also known as the "Jacobian".
d) Consider three independent uniform $(0, 1)$ variables. Compute the probability that the largest exceeds the sum of the other two.

Spring 2012, Applied Probability

1. 1st way: We can observe that

$$P(M=0) = \frac{1}{101} \quad \text{(prob. of the last person's place is empty)}$$

$$P(M=1) = \frac{1}{101} \quad \text{(prob of " the place of the person who is sitting at the place of the last one is empty ")}$$

In the same way we can say that

$$P(M=k) = \frac{1}{101} \quad \text{(prob. of "having the seat of } k^{th} \text{ person that is going to change place as empty")}$$

Thus, we have $P(M=k) = \frac{1}{101}$ For all $k=0,1,\dots,100$ Then,

$$E[M] = \sum_{k=0}^{100} k\,P(M=k) = \frac{1}{101}\sum_{k=1}^{100} k = \frac{1}{101} * \frac{100*101}{2} = 50$$

2nd way: Let M_n denote the number of "changes" when there are n seats and n tickets are sold. Then,

$$E[M_n] = E[M_n \mid n^{th} \text{ person's place is empty}]\,P(n^{th} \text{ person's place is empty})$$
$$+ E[M_n \mid n^{th} \text{ person's place is occupied}]\,P(n^{th} \text{ person's place is occupied})$$
$$= 0 + \left(E[M_{n-1}] + 1\right)\cdot\frac{n-1}{n}$$
$$= \frac{n-1}{n} + \frac{n-1}{n}E[M_{n-1}]$$

So,
$$E[M_{101}] = \frac{100}{101} + \frac{100}{101}E[M_{100}]$$
$$= \frac{100}{101} + \frac{99}{101} + \frac{99}{101}E[M_{99}]$$
$$\vdots$$
$$= \frac{100}{101} + \frac{99}{101} + \frac{98}{101} + \dots + \frac{1}{101} \qquad E[M_1] = 1$$
$$= \frac{1}{101}\sum_{k=1}^{100} k$$
$$= 50$$

2. a) $$E[S] = \sum_{1\le i<j\le n} E[1_{\{X_i = X_j\}}] = \sum_{1\le i<j\le n} P\{X_i = X_j\}$$

$$\rightarrow P\{X_i = X_j\} = \sum_{k=1}^{\infty} P\{X_i = X_j = k\} = \sum_{k=1}^{\infty} P\{X_i = k\}\,P\{X_j = k\} = \sum_{k=1}^{\infty} p_k^2 = f_2$$

So, $\mathbb{E}[S] = \sum_{1 \le i < j \le n} f_2 = \binom{n}{2} f_2$.

b) $\mathbb{E}[T] = \sum_{1 \le i < j < k \le n} \mathbb{E}[\mathbb{1}_{\{X_i = X_j = X_k\}}] = \sum_{1 \le i < j < k \le n} P\{X_i = X_j = X_k\}$

$$\rightarrow P\{X_i = X_j = X_k\} = \sum_{m=1}^{\infty} P\{X_i = X_j = X_k = m\} = \sum_{m=1}^{\infty} P\{X_i = m\} P\{X_j = m\} P\{X_k = m\}$$

$$= \sum_{m=1}^{\infty} P_m^3 = f_3$$

So, $E[T] = \sum_{1 \le i < j < k \le n} f_3 = \binom{n}{3} f_3$

c) $$S^2 = \sum_{1 \le i < j \le n} \mathbb{1}^2_{\{X_i = X_j\}} + 2 \sum_{1 \le i < j < k \le n} \mathbb{1}_{\{X_i = X_j\}} \mathbb{1}_{\{X_j = X_k\}} + 2 \sum_{1 \le i < j < k < l \le n} \mathbb{1}_{\{X_i = X_j\}} \mathbb{1}_{\{X_k = X_l\}}$$

$$E[S^2] = \sum_{1 \le i < j \le n} \mathbb{E}[\mathbb{1}_{\{X_i = X_j\}}] + 2 \sum_{1 \le i < j < k \le n} \mathbb{E}[\mathbb{1}_{\{X_i = X_j = X_k\}}] + 2 \sum_{1 \le i < j < k < l \le n} \mathbb{E}[\mathbb{1}_{\{X_i = X_j = X_k = X_l\}}]$$

$$= \binom{n}{2} f_2 + 2\binom{n}{3} f_3 + 2\binom{n}{4} f_4$$

$$\mathrm{Var}(S) = \binom{n}{2} f_2 + 2\binom{n}{3} f_3 + 2\binom{n}{4} f_4 - \binom{n}{2}^2 f_2^2$$

3. a) Firstly, we know that since $X \sim$ exponential(1), $X \sim$ gamma$(1,1)$ where the pdf of $Y \sim$ gamma(α, β) is

$$f_Y(y) = \frac{x^{\alpha-1} e^{-y/\beta}}{\Gamma(\alpha)\, \beta^{\alpha}}, \quad y > 0$$

Then, we know that if $Y_1 \sim$ gamma(α_1, β) and $Y_2 \sim$ gamma(α_2, β) are independent, then $Y_1 + Y_2 \sim$ gamma$(\alpha_1 + \alpha_2, \beta)$. Since X_1, X_2, X_3, X_4 are independent we have $S_4 = X_1 + X_2 + X_3 + X_4 \sim$ gamma$(4,1)$. So, S_4 has pdf,

$$f_{S_4}(t) = \frac{t^3 e^{-t}}{\Gamma(4)\, 1^4} = \frac{t^3 e^{-t}}{3!}, \quad t > 0$$

b) We know that for any iid r.v. $X_1, \dots, X_n$,

$$f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = \begin{cases} n!\, f_{X_1}(x_1) \cdots f_{X_n}(x_n), & -\infty < x_1 < \cdots < x_n < \infty \\ 0, & \text{otherwise} \end{cases}$$

Since $U_1, U_2, U_3 \sim$ uniform$(0,1)$, we know $f_{U_i}(u) = 1$, $0 < u < 1$ for $i = 1, 2, 3$.

So, we get the pdf $g(x,y,z)$ of $(U_{(1)}, U_{(2)}, U_{(3)})$ as

$$g(x,y,z) = \begin{cases} 3! & , 0<x<y,z<1 \\ 0 & , \text{otherwise} \end{cases}$$

c) We have X_1, X_2, X_3, X_4 and $S_n = \sum_{i=1}^{n} X_i$ for $n=1,2,3,4$. Now, let us find the pdf of S_1, S_2, S_3, S_4.

Noting $(J)_{ij} = \frac{\partial S_i}{\partial X_j}$, we have

$$J = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 1$$

Since $x_1, x_2, x_3, x_4 > 0$ we have $0 < s_1 < s_2 < s_4$ and, since $f_{X_1,X_2,X_3,X_4}(x_1,x_2,x_3,x_4) = e^{-(x_1+x_2+x_3+x_4)}$, we obtain that

$$f_{S_1,S_2,S_3,S_4}(s_1,s_2,s_3,s_4) = e^{-s_4} \quad , 0 < s_1 < s_2 < s_3 < s_4$$

Now, let $T_1 = \frac{S_1}{S_4}$, $T_2 = \frac{S_2}{S_4}$, $T_3 = \frac{S_3}{S_4}$ and $T_4 = S_4$ and let us find the pdf of T_1, T_2, T_3, T_4. Again noting that $(J) = \frac{\partial T_i}{\partial S_j}$, we have

$$J = \begin{vmatrix} 1/s_4 & 0 & 0 & -s_1/s_4^2 \\ 0 & 1/s_4 & 0 & -s_2/s_4^2 \\ 0 & 0 & 1/s_4 & -s_3/s_4^2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = \frac{1}{s_4^3}$$

Since $0<s_1<s_2<s_3<s_4$, we have $0 < t_1 t_4 < t_2 t_4 < t_3 t_4 < t_4$. Since $t_4 > 0$, we have $0 < t_1 < t_2 < t_3 < 1$ So,

$$f_{T_1,T_2,T_3,T_4}(t_1,t_2,t_3,t_4) = e^{-t_4} \cdot t_4^3 \quad , \quad 0 < t_1 < t_2 < t_3 < 1, \; t_4 > 0$$

$$= \frac{e^{-t_4} t_4^3}{6} * 6 , \quad 0 < t_1 < t_2 < t_3 < 1 , \; t_4 > 0$$

We know $f_{T_4}(t_4) = \frac{e^{-t_4} t_4^3}{6}$, $t_4 > 0$. So, we deduce that $(T_1, T_2, T_3) = \left(\frac{S_1}{S_4}, \frac{S_2}{S_4}, \frac{S_3}{S_4}\right)$ and $T_4 = S_4$ are independent and

$$f_{T_1,T_2,T_3}(t_1,t_2,t_3) = 6 , \quad 0 < t_1 < t_2 < t_3 < 1$$

Thus, we conclude that $\left(\frac{S_1}{S_4}, \frac{S_2}{S_4}, \frac{S_3}{S_4}\right) \overset{d}{=} (U_{(1)}, U_{(2)}, U_{(3)})$.

d) The desired probability can be expressed as

$$P(U_1 > U_2 + U_3) + P(U_2 > U_1 + U_3) + P(U_3 > U_1 + U_2)$$

and this probability is the same as $3P(U_1 > U_2 + U_3)$, by symmetry.

Now, let us find pdf of $U_2 + U_3$

$$f_{U_2+U_3}(t) = \int_0^1 f_{U_1}(x) f_{U_2}(t-x)\,dx = \int_0^1 f_{U_2}(t-x)\,dx$$

since $f_{U_2}(x) = 1$ when $0 < x < 1$. We know $0 < t < 2$ and $f_{U_2}(t-x) = 1$ only when $0 < t-x < 1 \Leftrightarrow t-1 < x < t$.

→ <u>If $0 < t < 1$:</u> $\quad f_{U_2+U_3}(t) = \int_0^t dx = t$

→ <u>If $1 \le t < 2$:</u> $\quad f_{U_2+U_3}(t) = \int_{t-1}^1 dx = 2-t$

So,

$$f_{U_1+U_2}(t) = \begin{cases} t & \text{if } 0 < t < 1 \\ 2-t & \text{if } 1 \le t < 2 \\ 0 & \text{otherwise} \end{cases}$$

Now, consider

$$P(U_1 > U_2 + U_3) = \int_0^1 \int_t^1 f_{U_1}(u) f_{U_2+U_3}(t)\,du\,dt$$

Since we have $t < u < 1$ and $0 < t < 1$, we must have $0 < t < u < 1$ and also $f_{U_1}(u) = 1$, $f_{U_2+U_3}(t) = t$ in the above integral So,

$$P(U_1 > U_2 + U_3) = \int_0^1 \int_t^1 t\,du\,dt = \int_0^1 (t - t^2)\,dt = \left[\frac{t^2}{2} - \frac{t^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

Thus, the desired probability is $3 * \frac{1}{6} = \frac{1}{2}$.

Applied Probability (505A) Graduate Exam

Fall 2011

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1. True or false: if A and B are events such that $0 < P(A) < 1$ and $P(B|A) = P(B|A^c)$, then A and B are independent. Justify your answer.

2. Suppose X and Y are independent, each with exponential density e^{-x} for $x > 0$, and let $Z = X - Y$.

(a) Calculate the density of Z.

(b) Calculate the moment generating function of X, the characteristic function of X, and the characteristic function of Z.

3. On the first day of class, the professor observed that there are m men and n women in the class. He says "I will bet even money, that some man-woman pair of students in this class have the same birthday."

Let W denote the number of man-woman pairs which have a birthday in common, so that the professor is betting that $W > 0$. (For example, if Fred, Bob, Mary, Jane, and Linda all have the same birthday, then $W \geq 6$.) You may assume that the birthdays of the $m + n$ students are distributed independently and uniformly over the 365 days of a non-leap year.

(a) Find an exact simple expression for the expectation of W.

(b) Find and simplify the variance of W.

(c) Suppose $m = 10$ and $n = 20$. Name a simple distribution that gives a good approximation for the distribution of W. (There may be more than one acceptable answer; no proof is required.)

(d) With $m = 10$ and $n = 20$, does the professor expect to win or lose money? You can give a heuristic approximation, but your answer should involve some calculation. Recall that $\ln 2 \approx 0.693$.

Fall 2011, Applied Probability :

1 Consider by definition,

$$P(B|A) = P(B|A^c) \Rightarrow \frac{P(A\cap B)}{P(A)} = \frac{P(A^c\cap B)}{P(A^c)}$$

$$\Rightarrow \frac{P(A\cap B)}{P(A)} = \frac{P(B) - P(A\cap B)}{1-P(A)}$$

$$\Rightarrow P(A\cap B) - \cancel{P(A)P(A\cap B)} = P(A)P(B) - \cancel{P(A)P(A\cap B)}$$

$$\Rightarrow P(A\cap B) = P(A)P(B)$$

Thus, the events A and B are independent.

2. a) $F_Z(z) = P(Z\leq z) = P(X-Y\leq z)$

→ If $z\geq 0$:

$$F_Z(z) = \int_0^\infty\int_0^{y+z} e^{-x}e^{-y}\,dx\,dy = \int_0^\infty e^{-y}\left[-e^{-x}\right]_0^{y+z}dy = \int_0^\infty e^{-y}\left(-e^{-y-z}+1\right)dy$$

$$= \int_0^\infty \left(e^{-y} - e^{-2y-z}\right)dy = \left[-e^{-y} + \frac{e^{-z}e^{-2y}}{2}\right]_0^\infty = 1 - \frac{e^{-z}}{2}$$

→ If $z<0$:

$$F_Z(z) = \int_{-z}^\infty\int_0^{y+z} e^{-x}e^{-y}\,dx\,dy = \int_{-z} e^{-y}\left[-e^{-x}\right]_0^{y+z}dy = \int e^{-y}\left(-e^{-y-z}+1\right)dy$$

$$= \int_{-z}^\infty \left(e^{-y} - e^{-2y-z}\right)dy = \left[-e^{-y} + \frac{e^{-z}\cdot e^{-2y}}{2}\right]_{-z}^\infty = e^{z} - \frac{e^{z}}{2} = \frac{e^{z}}{2}$$

So,

$$F_Z(z) = \begin{cases} \frac{e^z}{2} & \text{if } z<0 \\ 1-\frac{e^{-z}}{2} & \text{if } z\geq 0 \end{cases}$$

and thus,

$$f_Z(z) = \begin{cases} \frac{e^z}{2} & \text{if } z<0 \\ \frac{e^{-z}}{2} & \text{if } z\geq 0 \end{cases}$$

b) → $M_X(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx}e^{-x}\,dx = \int_0^\infty e^{x(t-1)}\,dx = \left.\frac{e^{x(t-1)}}{t-1}\right|_0^\infty$ * we must have $t-1<0$, i.e. $\underline{t<1}$

$$= \frac{1}{t-1}(0-1) = \frac{1}{1-t} \quad \text{where } t<1.$$

$$\rightarrow \varphi_X(t) = \mathbb{E}[e^{itX}] = \int_0^\infty e^{itx} e^{-x} dx = \int_0^\infty e^{x(it-1)} dx = \left[\frac{e^{x(it-1)}}{it-1}\right]_0^\infty$$

$$= \frac{1}{it-1}\left[e^{-x}\cos tx + i e^{-x}\sin tx\right]_0^\infty = \frac{1}{it-1}(-1) = \frac{1}{1-it}$$

$$\rightarrow \varphi_Z(t) = \mathbb{E}[e^{itZ}] = \int_{-\infty}^0 e^{itz}\cdot\frac{e^z}{2}dz + \int_0^\infty e^{itz}\cdot\frac{e^{-z}}{2}dz = \frac{1}{2}\int_{-\infty}^0 e^{z(it+1)}dz + \frac{1}{2}\int_0^\infty e^{z(it-1)}dz$$

$$= \frac{1}{2}\left[\frac{e^{z(it+1)}}{it+1}\right]_{-\infty}^0 + \frac{1}{2}\left[\frac{e^{z(it-1)}}{it-1}\right]_0^\infty = \frac{1}{2(it+1)} - \frac{1}{2(it-1)} = \frac{\cancel{it}-1-\cancel{it}-1}{2(-t^2-1)}$$

$$= \frac{1}{1+t^2}$$

3. Let us define $W = \sum_{i=1}^m \sum_{j=1}^n 1_{A_{ij}}$ where A_{ij} is the event that the birthday of i^{th} man and j^{th} women are the same (after numbering them).

a) $\mathbb{E}[W] = \sum_{i=1}^m \sum_{j=1}^n P(A_{ij})$

$$P(A_{ij}): \sum_{k=1}^{365} (\text{bday of } i^{th} \text{ man} = \text{bday of } j^{th} \text{ women} = k) = \sum_{k=1}^{365} \frac{1}{(365)^2} = \frac{1}{365}$$

So, $\mathbb{E}[W] = \sum_{i=1}^m \sum_{j=1}^n \frac{1}{365} = \frac{mn}{365}$

b) Define $R_i := \sum_{k=1}^n 1_{A_{ik}}$. So, $W = \sum_{i=1}^m R_i$ and

$$Var(W) = \sum_{i=1}^m Var(R_i) + 2\sum_{1\le i<j\le m} Cov(R_i, R_j)$$

$$\rightarrow Var(R_i) = \sum_{k=1}^n Var(1_{A_{ik}}) + 2\sum_{1\le k<\ell\le n} Cov(1_{A_{ik}}, 1_{A_{i\ell}}).$$

$$\rightarrow Var(1_{A_{ik}}) = P(A_{ik}) - P(A_{ik})^2 = \frac{1}{365}\left(1-\frac{1}{365}\right) = \frac{364}{(365)^2}$$

$$\rightarrow Cov(1_{A_{ik}}, 1_{A_{i\ell}}) = P(A_{ik}\cap A_{i\ell}) - P(A_{ik})P(A_{i\ell}) = \frac{1}{(365)^2} - \frac{1}{(365)^2} = 0, \quad \forall k,\ell.$$

$$\Rightarrow Var(R_i) = \frac{364n}{(365)^2}$$

$$\rightarrow Cov(R_i, R_j) = Cov\left(\sum_{k=1}^n 1_{A_{ik}}, \sum_{\ell=1}^n 1_{A_{j\ell}}\right) = \sum_{k=1}^n \sum_{\ell=1}^n Cov(1_{A_{ik}}, 1_{A_{j\ell}})$$

$$\rightarrow \text{If } k=\ell: \quad Cov(1_{A_{ik}}, 1_{A_{j\ell}}) = P(A_{ik}\cap A_{j\ell}) - P(A_{ik})P(A_{j\ell}) = \frac{1}{(365)^2} - \frac{1}{(365)^2} = 0$$

$\rightarrow$ If $k \neq l$: $Cov(1_{A_{ik}}, 1_{A_{jl}}) = P(A_{ik} \cap A_{jl}) - P(A_{ik})P(A_{jl}) = \frac{1}{(365)^3} - \frac{1}{(365)^2}$

$$= \frac{1}{(365)^2} \cdot \frac{-364}{365} = \frac{-364}{(365)^3}$$

$$\Rightarrow Cov(R_i, R_j) = \frac{-364\, n(n-1)}{(365)^3}$$

$$\Rightarrow Var(W) = \sum_{i=1}^{m} \frac{364n}{(365)^2} + 2 \sum_{1 \le i < j \le m} \frac{-364n(n-1)}{(365)^3}$$

$$= \frac{364mn}{(365)^2} - \frac{364\, mn(m-1)(n-1)}{(365)^3}$$

$$= \frac{364mn}{(365)^2}\left[1 - \frac{(m-1)(n-1)}{365}\right]$$

$$= \frac{364mn\,(365 - (m-1)(n-1))}{(365)^3}$$

c) There are mn pairs to consider and if a pair has the same birthday, which has probability $1/365$, we can think this event to be success. We know that pairs are not independent, but approximately, we can say that $W \sim Binomial(mn, 1/365)$. Since $m = 10$, $n = 20$, we have $W \sim Binomial(200, 1/365)$. Then, we have $mn = 200 \geq 100$ and $p = \frac{1}{365} \leq 0.01$. So, actually, we can use Poisson approximation to Binomial distribution. So, we can say that $W \sim Poisson\left(\frac{200}{365}\right)$ where $\frac{200}{365} \approx 0.55$.

d) We want to find an approximate value for $P(W>0)$ or a bound for this probability. Consider

$$P(W>0) = 1 - P(W=0) = 1 - e^{-0.55} \cdot \frac{(0.55)^0}{0!} = 1 - e^{-0.55} < 1 - e^{-0.693} = 1 - e^{-\ln 2} = \frac{1}{2}$$

Thus, we have $P(W>0) < \frac{1}{2}$ showing that probability for the professor to win is less than $1/2$. So, in this case the professor should expect to lose money.

3 correct 7

4

MATH 505a QUALIFYING EXAM September 21, 2010

Solve all four problems. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

Problem 1. Let X_k, $k \geq 1$, be iid random variables with mean 1 and variance 1. Show that the limit

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k}{\sum_{k=1}^{n} X_k^2}$$

exists in an appropriate sense, and identify the limit.

Problem 2. Fix $p \in (0,1)$ and consider independent Poisson random variables X_k, $k \geq 1$ with

$$\mathbb{E}X_k = \frac{p^k}{k}.$$

Verify that the sum $\sum_{k=1}^{\infty} kX_k$ converges with probability one and determine the distribution of the random variable $Y = \sum_{k=1}^{\infty} kX_k$.

SUGGESTION: compute the generating function for X_k, for kX_k, and for Y.

Problem 3. A coin-making machine produces quarters in such a way that, for each coin, the probability U to turn up heads is uniform on $[0,1]$. A coin pops out of the machine.

(a) Compute the conditional distribution of U given that the coin is flipped once and lands on head.

(b) Compute, either exactly or approximately, the conditional distribution of U given that the coin is flipped 2000 times and lands on head 1500 times.

Problem 4. An ordered vertical stack of n books is on my desk. Every day, I pick one book uniformly at random from the stack and put the book on top of the stack. What is the expected number of days before the books are back to the original order?

COMMENTS: (a) For partial credit, just guess the answer, as a function of n. (b) For more credit, give a heuristic justification. (c) For bonus credit, give a proof.

Fall 2010, Applied Probability:

1. Let $Y_n =: \frac{1}{n}\sum_{k=1}^{n} X_k$ and $Z_n = \frac{1}{n}\sum_{k=1}^{n} X_k^2$. Since X_k's are iid, by Law of Large Numbers, $Y_n \xrightarrow{P} E[X_k] = 1$ as $n \to \infty$. Similarly, $Z_n \xrightarrow{P} E[X_k^2] = Var(X_k) + E[X_k]^2 = 2$ as $n \to \infty$. Then, since $g = 1/x$ is continuous a.e. we can say that $g(Z_n) = \frac{1}{\frac{1}{n}\sum_{k=1}^{n} X_k^2} \xrightarrow{P} \frac{1}{2}$ as $n \to \infty$

Finally, we obtain that

$$\frac{\sum_{k=1}^{n} X_k}{\sum_{k=1}^{n} X_k^2} = \left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \cdot \frac{1}{\frac{1}{n}\sum_{k=1}^{n} X_k^2} \xrightarrow{P} 1 \cdot \frac{1}{2} \quad \text{as } n \to \infty$$

2. We have $E[kX_k] = p^k$ and $\sum_{k=1}^{\infty} p^k = \frac{p}{1-p} < \infty$, and $Var(kX_k) = kp^k$ and $\sum_{k=1}^{\infty} kp^k = \frac{p}{(1-p)^2} < \infty$ Since the sequence $\{Y_k\}_{k\geq 1}$ is independent, by Kolmogorov's Two Series Theorem, we conclude that $\sum_{k=1}^{\infty} kX_k$ converges a.s.

Now, let us find the distribution of the r.v. Y. Let X be a Poisson(λ) r.v. and consider

$$E[s^X] = \sum_{k=0}^{\infty} s^k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}$$

So, $E[s^{X_k}] = e^{\frac{p^k}{k}(s-1)}$ and $E[s^{kX_k}] = E[(s^k)^{X_k}] = e^{\frac{p^k}{k}(s^k-1)}$.

Then, define $Y_n = \sum_{k=1}^{n} kX_k$ and consider

$$E[s^{Y_n}] = E\left[\prod_{k=1}^{n} s^{kX_k}\right] = \prod_{k=1}^{n} e^{\frac{p^k}{k}(s^k-1)} = e^{\sum_{k=1}^{n} \frac{(ps)^k}{k} - \sum_{k=1}^{n} \frac{p^k}{k}}$$

Now, by taking limit as $n \to \infty$ (since we are allowed to take limit inside of expectation here), we get

$$E[s^Y] = e^{\sum_{k=1}^{\infty} \frac{(ps)^k}{k} - \sum_{k=1}^{\infty} \frac{p^k}{k}} = e^{-\ln(1-ps) + \ln(1-p)}, \quad -1 \leq ps < 1 \text{ and } -1 \leq p < 1$$

$$= \frac{1-p}{1-sp}, \quad -1 \leq sp < 1$$

Since we know $P\{X=k\} = \frac{G_X^{(k)}(0)}{k!}$, denoting the probability generating function of Y by $G(s)$, observe that

$$G(s) = \frac{1-p}{1-sp}, \quad G'(s) = (1-p)\frac{p}{(1-sp)^2}, \quad G''(s) = (1-p)\frac{2p^2}{(1-sp)^3},$$

$$G'''(s) = (1-p)\frac{3.2p^3}{(1-sp)^3}, \quad \dots, \quad G^{(k)}(s) = (1-p)\frac{k!\,p^k}{(1-sp)^k}, \quad \text{for } k=0,1,2,\dots$$

So, we get

$$P\{Y=k\} = (1-p)p^k, \quad k=0,1,2,\dots$$

which gives us the distribution of Y.

3. a) Define

$$A = \begin{cases} 1 & \text{if the coin is flipped once and lands on head} \\ 0 & \text{otherwise} \end{cases}$$

We want to find $f_{U|A}(u|1)$, where $U \sim \text{Uniform}[0,1]$.

$$f_{U|A}(u|1) = \frac{f_{U,A}(u,1)}{f_A(1)} = \frac{f_{A|U}(1|u)\, f_U(u)}{f_A(1)} = \frac{u \cdot 1}{f_A(1)}$$

$$\rightarrow f_A(1) = \int_0^1 f_{A|U}(1|u)\, f_U(u)\,du = \int_0^1 u \cdot 1\,du = \frac{u^2}{2}\Big]_0^1 = \frac{1}{2}$$

So, we obtain that

$$f_{U|A}(u|1) = 2u, \quad 0 \leq u \leq 1,$$

as the desired distribution.

b) In a similar way in part (a), define

$$B = \begin{cases} 1 & \text{if the coin is flipped 2000 times and lands on head 1500 times} \\ 0 & \text{otherwise} \end{cases}$$

Now, we want to find $f_{U|B}(u,1)$, where $U \sim \text{Uniform}[0,1]$

$$f_{U|B}(u|1) = \frac{f_{B|U}(1|u)\, f_U(u)}{f_B(1)} = \frac{\binom{2000}{1500} u^{1500} (1-u)^{500} \cdot 1}{f_B(1)}$$

$$f_B(1) = \int_0^1 f_{B|U}(1|u)\, f_U(u)\, du = \int_0^1 \binom{2000}{1500} u^{1500} (1-u)^{500}\, du$$

Then,

$$f_{U|B}(u|1) = \frac{\binom{2000}{1500} u^{1500} (1-u)^{500}}{\int_0^1 \binom{2000}{1500} u^{1500} (1-u)^{500}\, du} = \frac{u^{1500}\ u^{500}}{\int_0^1 u^{1500} (1-u)^{500}\, du} \quad , 0<u<1$$

2, pgf of τ_n

MATH 505a **Screening Exam** **Spring, 2010**

Last Name: ____________________ **First Name:**____________________

1. Let (X_1, X_2) be standard bivariate normal random variables with correlation $\rho = \frac{3}{5}$. Let (Y_1, Y_2) denote the midterm exam score and final exam score of a randomly selected student in class. Assume

$$Y_1 = 80 + 3X_1, Y_2 = 75 + 2X_2.$$

If a student got 90 in the midterm exam,

a) what is the conditional expectation and conditional variance of his/her final exam score?

b) What is the conditional probability that he/she got more than 75 in the final exam?

2. Let $\{S_k\}_{k\geq 0}, S_0 = 0$, be a symmetric simple random walk. For an integer $n \geq 1$, let $\tau_n = \min\{k \geq 1 : S_k \notin (-n, n)\}$ be the first time k such that S_k leaves the region $(-n, n), n \geq 1$, where $\tau_n = \infty$ if there is no such k. Find the moment generating function of $S_{\tau_n}, \mathbf{E}S_{\tau_n}$ and $\text{Var}(S_{\tau_n})$.

3. Let $\Theta_1, \Theta_2, \ldots$ be a sequence of independent, identically distributed random variables with the uniform distribution on the interval $(0, 2\pi)$. For $n = 1, 2, \ldots$ define

$$X_n = \sum_{k=1}^{n} \cos\Theta_k, Y_n = \sum_{k=1}^{n} \sin\Theta_k, \text{ and } R_n^2 = X_n^2 + Y_n^2.$$

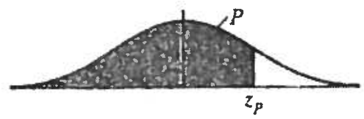
Show that

a) there is a sequence of numbers a_n so that $(a_n X_n, a_n Y_n)$ has a limiting bivariate normal distribution as $n \to \infty$;

b) $\lim_{n\to\infty} \mathbf{P}(R_n^2 \geq n)$ exists.

TABLE 2

Cumulative Normal Distribution—Values of P Corresponding to z_p for the Normal Curve



z is the standard normal variable. The value of P for $-z_p$ equals 1 minus the value of P for $+z_p$; for example, the P for -1.62 equals $1 - .9474 = .0526$.

z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Spring 2010, Applied Probability.

1 Since (X_1, X_2) has standard bivariate distribution, we know that given $X_1 = x$, X_2 has $N\left(0+\frac{3}{5}\cdot\frac{1}{1}(x-0), \left(1-\left(\frac{3}{5}\right)^2\right)1\right) = N\left(\frac{3x}{5}, \frac{16}{25}\right)$ distribution.

a) $E[Y_2 | Y_1 = 90] = E[75+2X_2 | 80+3X_1 = 90] = 75 + 2E[X_2 | X_1 = \frac{10}{3}]$. Since, given $X_1 = \frac{10}{3}$, $X_2 \sim N\left(2, \frac{16}{25}\right)$, we deduce that $E[Y_2 | Y_1 = 90] = 79$.

b) $P(Y_2 > 75 | Y_1 = 90) = P(75+2X_2 > 75 | 80+3X_1 = 90)$

$$= P(X_2 > 0 | X_1 = \tfrac{10}{3})$$

$$= P(X > 0) \quad \text{where} \quad X \sim N\left(2, \tfrac{16}{25}\right)$$

$$= P\left(\frac{X-2}{4/5} > -\frac{2}{4/5}\right)$$

$$= P(Z > -2.5)$$

$$\approx 0.9938$$

2. $$E[e^{tS_{\tau_n}} | \tau_n = k] = E[e^{tS_k} | \tau_n = k]$$

$$= E\left[e^{t\sum_{i=1}^{k} X_i} | \tau_n = k\right]$$

by independence

$$= \prod_{i=1}^{k} E[e^{tX_i}]$$

$$= \left(\frac{e^t + e^{-t}}{2}\right)^k$$

So, $E[e^{tS_{\tau_n}} | \tau_n] = \left(\frac{e^t+e^{-t}}{2}\right)^{\tau_n}$. Now, we need to find the probability generating function of τ_n.

Let us call the probability generating function of τ_n by $G(s)$

So,

$$M_{S_{\tau_n}}(t) = G\left(\frac{e^t + e^{-t}}{2}\right) = G(\cosh t)$$

$$M'_{S_{\tau_n}}(t) = G'(\cosh t)\sinh t$$

$\rightarrow$ $E[S_{\tau_n}] = M'_{S_{\tau_n}}(0) = G'(1) * 0 = 0$

$$M''_{S_{\tau_n}}(t) = G''(\cosh t)(\sinh t)^2 + G'(\cosh t)\cosh t$$

$\rightarrow$ $E[S^2_{\tau_n}] = M''_{S_{\tau_n}}(0) = G'(1) * 1 = E[\tau_n]$.

$\rightarrow$ $Var(S_{\tau_n}) = E[S^2_{\tau_n}] - (E[S_{\tau_n}])^2 = E[\tau_n]$

3. a) Suppose we have the random sample $\begin{bmatrix} \cos\theta_1 \\ \sin\theta_1 \end{bmatrix}, \begin{bmatrix} \cos\theta_2 \\ \sin\theta_2 \end{bmatrix}, \dots, \begin{bmatrix} \cos\theta_k \\ \sin\theta_k \end{bmatrix}, \dots$

$\rightarrow$ $E[\cos\theta_i] = \int_0^{2\pi} \cos\theta \frac{1}{2\pi} d\theta = 0$, $E[\sin\theta_i] = \int_0^{2\pi} \sin\theta \frac{1}{2\pi} d\theta = 0$

$\rightarrow$ $E[\cos^2\theta_i] = \int_0^{2\pi} \cos^2\theta \frac{1}{2\pi} d\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{1+\cos 2\theta}{2} d\theta = \frac{1}{4\pi}\left[\theta + \frac{\sin 2\theta}{2}\right]_0^{2\pi} = \frac{1}{2}$

$\rightarrow$ $E[\sin^2\theta_i] = \int_0^{2\pi} \sin^2\theta \frac{1}{2\pi} d\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-\cos 2\theta}{2} d\theta = \frac{1}{4\pi}\left[\theta - \frac{\sin 2\theta}{2}\right]_0^{2\pi} = \frac{1}{2}$

$\rightarrow$ $Var(\cos\theta_i) = \frac{1}{2}$, $Var(\sin\theta_i) = \frac{1}{2}$

$\rightarrow$ $E[\cos\theta_i \sin\theta_i] = \int_0^{2\pi} \cos\theta\sin\theta \frac{1}{2\pi} d\theta = \frac{1}{2\pi}\left[\frac{\sin^2\theta}{2}\right]_0^{2\pi} = 0$

Now, let us choose $a_n = \frac{\sqrt{2}}{n}$ and define $A_k = \sqrt{2}\cos\theta_k$ and $B_k = \sqrt{2}\sin\theta_k$ Then, we get

$E[A_k] = E[B_k] = 0$, $Var(A_k) = Var(B_k) = 1$, $Cov(A_k, B_k) = 0$

Then, by using multivariate CLT,

$$\sqrt{n}\begin{bmatrix} a_n X_n \\ b_n Y_n \end{bmatrix} = \sqrt{n}\begin{bmatrix} \frac{\sqrt{2}}{n} X_n \\ \frac{\sqrt{2}}{n} Y_n \end{bmatrix} = \sqrt{n}\begin{bmatrix} \frac{1}{n}\sum_{k=1}^{n} A_k \\ \frac{1}{n}\sum_{k=1}^{n} B_k \end{bmatrix} \xrightarrow{d} N(0, I)$$

Thus, if we choose $a_n = \frac{\sqrt{2}}{n}$, $(a_n X_n, a_n Y_n)$ has limiting bivariate normal distribution ...

b) By part (a), we know that $\sqrt{\frac{2}{n}}X_n \xrightarrow{d} N(0,1)$ and $\sqrt{\frac{2}{n}}Y_n \xrightarrow{d} N(0,1)$ as $n \to \infty$ Since $g(x)=x^2$ is a continuous function, $\frac{2}{n}X_n^2 \xrightarrow{d} \chi^2_1$ and $\frac{2}{n}Y_n^2 \xrightarrow{d} \chi^2_1$ as $n \to \infty$ Then by independence, $\frac{2}{n}(X_n^2+Y_n^2) \xrightarrow{d} \chi^2_2$ as $n \to \infty$ We know $\chi^2_2 = \Gamma(1,2) = \text{exponential}(2)$

So, $\frac{2}{n}(X_n^2+Y_n^2) \xrightarrow{d} T$ where $T \sim \text{exponential}(2)$

Now, observe that

$$P(R_n^2 \geq n) = P(X_n^2+Y_n^2 \geq n) = P\left(\frac{2}{n}(X_n^2+Y_n^2) \geq 2\right) = 1-P\left(\frac{2}{n}(X_n^2+Y_n^2)<2\right)$$

$$\Rightarrow \lim_{n\to\infty} P(R_n^2 \geq n) = 1-\lim_{n\to\infty} P\left(\frac{2}{n}(X_n^2+Y_n^2)<2\right) = 1-P(T<2) = 1-\int_0^2 \frac{1}{2}e^{-x/2}dx$$

$$= 1-\left[-e^{-x/2}\right]_0^2 = 1-(-e^{-1}+1) = e^{-1}$$

Note for question #2: Define $A_k = n-S_k$ and $B_k = n+S_k$. Then

$$\tau_n = \inf\{k \geq 1 \; S_k \notin (-n,n)\} = \inf\{k \geq 1 \; A_k B_k = 0\}$$

Now, we want to find $E[\tau_n]$. Firstly consider that A_nB_n+n is a martingale

$$E[A_{n+1}B_{n+1}+n+1 \mid A_n, B_n] = \frac{(A_n+1)(B_n-1)}{2} + \frac{(A_n-1)(B_n+1)}{2} + n+1$$

$$= \frac{A_nB_n - \cancel{A_n} + \cancel{B_n} - 1 + A_nB_n + \cancel{A_n} - \cancel{B_n} - 1}{2} + n+1$$

$$= \frac{2(A_nB_n-1)}{2} + n+1$$

$$= A_nB_n + n$$

So, by Optional Stopping Theorem, $E[\tau_n] = E[A_{\tau_n}B_{\tau_n}+\tau_n] = E[A_1B_1+1]$

$\to E[A_1B_1+1] = n^2+nE[S_1]-nE[S_1]-E[S_1^2]+1 = n^2-E[X_1^2]+1 = n^2-1+1 = n^2$

Thus $E[\tau_n]=n^2$ and $Var(S_{\tau_n}) = n^2$.

MATH 505a Screening Exam Fall, 2009

Last Name: __________________ **First Name:**__________________

1. A drawer contains N pairs of socks; each sock has *precisely* one mate. The $2N$ socks are randomly arranged in the drawer. I choose k socks (randomly) from among the $2N$ socks in the drawer, with $2 \le k \le 2N$. What is the *expected* number of complete pairs in my sample of k socks?

2. A clerk in a gas station is rolling a fair dice while waiting for the customers to come. Suppose that the number of times the dice is rolled between two customers has a Poisson distribution with parameter $\lambda = 5$. Let ξ be the total points (of the dice) the clerk observed right before the next customer comes in. Determine $E\xi$ and $D\xi$ (standard deviation).

3. Let a random variable X be normal $N(\mu, \sigma^2)$ and let the conditional distribution of Y given X be normal $N(a_1 + a_2X, \sigma_1^2)$.

(a) Find the joint probability density function of X and Y.

(b) Find the marginal distribution of Y and the correlation coefficient of X and Y.

4. Let ξ and η be two random variables, both taking only two values. Show that if they are uncorrelated, then they are independent.

Fall 2009, Applied Probability:

1. Let us number chosen socks from 1 to k and A_{ij} denote the event that i^{th} and j^{th} ones are the same, $1 \le i < j \le k$. Then M denote the number of pairs among k socks. So, we can write

$$M = \sum_{1 \le i < j \le k} 1_{A_{ij}}$$

and so,

$$E[M] = \sum_{1 \le i < j \le k} P(A_{ij})$$

For any $1 \le i < j \le k$, $P(A_{ij}) = N * \dfrac{\binom{2}{2}\binom{2N-2}{0}}{\binom{2N}{2}} = \dfrac{2N}{2N(2N-1)} = \dfrac{1}{2N-1}$

Thus, $E[M] = \sum_{1 \le i < j \le k} \dfrac{1}{2N-1} = \binom{k}{2} \dfrac{1}{2N-1} = \dfrac{k(k-1)}{2(2N-1)}$.

2. Let ξ_n be the total number of dots that are observed in n throws. Then letting $\xi_n^{(i)}$ to be the number of dots that are observed at i^{th} throw, we can write $\xi_n = \xi_n^{(1)} + \dots + \xi_n^{(n)}$. Now, consider

$$E[\xi \mid N=n] = E[\xi_n \mid N=n] = \sum_{i=1}^{n} E[\xi_n^{(i)} \mid N=n] = \sum_{i=1}^{n} \left(\sum_{k=1}^{6} k P\{\xi_n^{(i)} = k \mid N=n\} \right)$$

$$= \sum_{i=1}^{n} \left(\sum_{k=1}^{6} k P\{\xi_n^{(i)} = k\} \right) = \sum_{i=1}^{n} \left(\sum_{k=1}^{6} \frac{k}{6} \right) = \sum_{i=1}^{n} \frac{7}{2} = \frac{7n}{2}$$

Thus $E[\xi \mid N] = \dfrac{7N}{2}$ and so, $E[\xi] = E\left[E[\xi \mid N]\right] = E\left[\dfrac{7N}{2}\right] = \dfrac{7}{2} E[N] = \dfrac{35}{2}$

Now, let us find $E[\xi^2 \mid N]$. Consider

$$E[\xi^2 \mid N=n] = E[\xi_n^2 \mid N=n] = \sum_{i=1}^{n} E[\xi_n^{(i)2} \mid N=n] + 2 \sum_{1 \le i < j \le n} E[\xi_n^{(i)} \xi_n^{(j)} \mid N=n]$$

$$= \sum_{i=1}^{n} \left(\sum_{k=1}^{6} k^2 P\{\xi_n^{(i)} = k\} \right) + 2 \sum_{1 \le i < j \le n} \left(\sum_{k=1}^{6} \sum_{\ell=1}^{6} k\ell P\{\xi_n^{(i)} = k, \xi_n^{(j)} = \ell\} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{1}{6} \sum_{k=1}^{6} k^2 \right) + 2 \sum_{1 \le i < j \le n} \left(\frac{1}{36} \sum_{k=1}^{6} \left(k \sum_{\ell=1}^{6} \ell \right) \right)$$

$$= \sum_{i=1}^{n} \frac{91}{6} + 2 \sum_{1 \le i < j \le n} \frac{49}{4}$$

$$= \frac{91n}{6} + \frac{49n(n-1)}{4}$$

$$= \frac{182n + 147n(n-1)}{12}$$

$$= \frac{147n^2 + 35n}{12}$$

So, $E[\xi^2 | N] = \dfrac{147N^2 + 35N}{12}$ and $Var(\xi | N) = \dfrac{147N^2 + 35N}{12} - \dfrac{49N^2}{4} = \dfrac{35N}{12}$

Thus,

$$Var(\xi) = Var(\mathbb{E}[\xi | N]) + \mathbb{E}[Var(\xi | N)]$$

$$= Var\left(\frac{7N}{2}\right) + \mathbb{E}\left[\frac{35N}{12}\right]$$

$$= \frac{49}{4} Var(N) + \frac{35}{12} E[N]$$

$$= 5 * \frac{147 + 35}{12}$$

$$= \frac{935}{12}$$

and $D\xi = \dfrac{\sqrt{935}}{2\sqrt{3}}$

3. a) $X \sim N(\mu, \sigma^2)$ and $Y|X = N(a_1 + a_2 X, \sigma_1^2)$

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) f_X(x)$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left\{-\frac{1}{2\sigma_1^2}(y - a_1 - a_2 x)^2\right\} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

$$= \frac{1}{2\pi \sigma_1 \sigma} \exp\left\{-\frac{1}{2\sigma_1^2}(y - a_1 - a_2 x)^2 - \frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

b) Since we have $Y|X \sim N(a_1 + a_2 X, \sigma_1^2)$, we have

$Y|X - (a_1 + a_2 X) \sim N(0, \sigma_1^2) \Rightarrow Y|X - (a_1 + a_2 X)|X \sim N(0, \sigma_1^2)$

$\Rightarrow (Y - a_1 - a_2 X) | X \sim N(0, \sigma_1^2) \Rightarrow Y - a_1 - a_2 X \perp\!\!\!\perp X$ and $Y - a_1 - a_2 X \sim N(0, \sigma_1^2)$.

Since $a_1 + a_2 X \sim N(a_1 + a_2 \mu, a_2^2 \sigma^2)$, we obtain

$$Y = (Y - a_1 - a_2 X) + a_1 + a_2 X \sim N(a_1 + a_2\mu,\ \sigma_1^2 + a_2^2\sigma^2)$$

Thus Y has normal distribution with mean $a_1 + a_2\mu$ and varionce $\sigma_1^2 + a_2^2\sigma^2$.

Since $Y - a_1 - a_2 X$ and X are independent, $Cov(Y - a_1 - a_2 X, X) = 0$. so,

$$0 = Cov(Y, X) - a_2 Var(X) \Rightarrow Cov(X, Y) = a_2\sigma^2.$$

Thus, the correlation coefficient ρ of X and Y is

$$\rho = \frac{a_2\sigma^2}{\sigma\sqrt{\sigma_1^2 + a_2^2\sigma^2}} = \frac{a_2\sigma}{\sqrt{\sigma_1^2 + a_2^2\sigma^2}}$$

4. This is a part of question 1 in Fall 2014, and solved there

MATH 505a | QUALIFYING EXAM | February 5, 2009

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1.) Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. random variables with a continuous distribution function. Let

$$N = \min\{n \geq 2\colon\ X_n > X_{n-1}\}.$$

a) Find $\mathbb{P}(N = n)$ for $n = 2, 3, \ldots$.
b) Find $\mathbb{P}(N \geq n)$ for $n = 1, 2, 3, \ldots$.
c) Find and simplify EN.

2.) A deck (deck #1) of cards has a red cards and b black cards, another deck (deck #2) has c red cards and d black cards. Assume $a, b, c, d \geq 1$. Both decks are well-shuffled. Suppose that you pick f cards ($0 \leq f \leq a + b$) randomly from deck #1 and mix them into deck #2. One card is now selected and removed from the mixed up deck #2.
a) What is the chance that the first card selected is red?
b) Given that the first selected card is red, what is the chance that it originally came from deck #1 ?
c) Are the two events, that the first selected card is red, and that the first selected card originally came from deck #1, independent? [Possible answers include YES, NO, or some relation among the parameters a, b, c, d, f. If you give a relation, please simplify it.]

3.) Suppose that the random variables Y, X, X_1, X_2, ..., are independent, identically distributed, and that $P\{Y = n\} = 2^{-n}$, for $n = 1, 2, \ldots$; and $P\{X \geq t\} = e^{-\pi t}$, for $t > 0$ and $k = 1, 2, \ldots$. Let $S_n = X_1 + X_2 + \cdots + X_n$. Let $Z = S_Y$.

a) Calculate and simplify $\mathbb{E}\,Z$.
b) Simplify the probability generating function $\mathbb{E}\,s^Y$.
c) Simplify the moment generating function $\mathbb{E}\,\exp(\beta X)$.
d) Simplify $\mathbb{E}\,\exp(\beta Z)$.

e) Calculate and simplify $\mathbb{E}\,Z^3$.

[Hint, if you answered a) by some easy method, use this to check your results for d) and e).]

Spring 2009, Applied Probability:

1. a) $P(N=n) = P\{X_1 \geq X_2 \geq X_3 \geq \dots \geq X_{n-2} \geq X_{n-1},\ X_{n-1} < X_n\}$ $\Big\}$ since X_i's are continuous r.v.

$= P\{X_1 > X_2 > X_3 > \dots > X_{n-2} > X_{n-1},\ X_{n-1} < X_n\}$

Since X_i's are iid, any ordering of X_i's is equally likely. That is the events $\{X_{\pi(1)} > X_{\pi(2)} > \dots > X_{\pi(n)}\}$ are all have the same probability for all permutations Since there are $n!$ possibilities, any ordering has probability $\frac{1}{n!}$. On the other hand for the event $\{X_1 > X_2 > \dots > X_{n-1},\ X_{n-1} < X_n\}$, there are $n-1$ possible places for X_n. Thus, we deduce,

$$P(N=n) = P\{X_1 > X_2 > \dots > X_{n-1},\ X_{n-1} < X_n\} = \frac{n-1}{n!}$$

b) $P(N \geq n) = P\{X_1 > X_2 > \dots > X_{n-1}\} = \dfrac{1}{(n-1)!}$

c) $E[N] = \sum_{n=2}^{\infty} n P\{N=n\} = \sum_{n=2}^{\infty} n \cdot \frac{n-1}{n!} = \sum_{n=2}^{\infty} \frac{1}{(n-2)!} = e$

2. a) Let A be the event that the first chosen card is red, and let B denote the number of red cards among f cards. Then,

$$P(A) = \sum_{k=0}^{f} P(A \mid B=k) P(B=k)$$

$$= \sum_{k=0}^{f} \frac{c+k}{c+d+f} \frac{\binom{a}{k}\binom{b}{f-k}}{\binom{a+b}{f}}, \quad f-b \leq k \leq a$$

$$= \frac{c}{c+d+f} \underbrace{\sum_{k=0}^{f} \frac{\binom{a}{k}\binom{b}{f-k}}{\binom{a+b}{f}}}_{=1} + \frac{1}{c+d+f} \underbrace{\sum_{k=0}^{f} k \frac{\binom{a}{k}\binom{b}{f-k}}{\binom{a+b}{f}}}_{\substack{= E[X] \text{ where } X \sim \text{Hypergeometric}(a+b,a,f) \\ = \frac{fa}{a+b}}}$$

$$= \frac{1}{c+d+f}\left(c + \frac{af}{a+b}\right)$$

$$= \frac{ac+bc+af}{(a+b)(c+d+f)}$$

b) Let us denote the event that card comes from deck #1 by C. So, we are trying to find $P(C|A) = \dfrac{P(C\cap A)}{P(A)}$

$$P(C\cap A) = \sum_{k=0}^{f} P(C\cap A \mid B=k)\,P(B=k)$$

$$= \sum_{k=0}^{f} \frac{k}{c+d+f} \frac{\binom{a}{k}\binom{b}{f-k}}{\binom{a+b}{f}}, \quad f-b \le k \le a$$

$$= \frac{1}{c+d+f} \cdot \frac{af}{a+b}$$

$$= \frac{af}{(c+d+f)(a+b)}$$

So, $P(C|A) = \dfrac{af}{(c+d+f)(a+b)} \, \dfrac{(a+b)(c+d+f)}{ac+bc+af} = \dfrac{af}{ac+bc+af}$

c) Now, consider

$$P(C) = \sum_{k=0}^{f} P(C \mid B=k)\,P(B=k)$$

$$= \sum_{k=0}^{f} \frac{f}{c+d+f} \frac{\binom{a}{k}\binom{b}{f-k}}{\binom{a+b}{f}}, \quad f-b\le k\le a$$

$$= \frac{f}{c+d+f}$$

We see that $P(A)P(C) = \dfrac{(ac+bc+af)f}{(a+b)(c+d+f)^2}$, $P(A\cap C) = \dfrac{af}{(c+d+f)(a+b)}$

$$P(A)P(C) = P(A\cap C) \iff \frac{ac+bc+af}{(c+d+f)} = a \iff ac+bc+af = ac+ad+af$$

$$\iff bc = ad$$

So, if $bc = ad$, A and C are independent

3. a) $\mathbb{E}[Z] = \mathbb{E}\big[\mathbb{E}[S_Y|Y]\big]$

$$\mathbb{E}[S_Y|Y=n] = \sum_{i=1}^{n} \mathbb{E}[X_i|Y=n] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n}\left(\int_0^\infty \pi t e^{-\pi t}\,dt\right) = \sum_{i=1}^{n} \frac{1}{\pi} = \frac{n}{\pi}$$

Thus $\mathbb{E}[S_Y|Y] = \dfrac{Y}{\pi}$

$$\mathbb{E}[Z] = \frac{1}{\pi}\mathbb{E}[Y] = \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{k}{2^k} = \frac{2}{\pi}$$ by the note below.

b) $$\mathbb{E}[s^Y] = \sum_{n=1}^{\infty} s^n P\{Y=n\} = \sum_{n=1}^{\infty}\left(\frac{s}{2}\right)^n = \frac{s}{2-s}, \quad |s|<2$$

c) $$\mathbb{E}[e^{\beta X}] = \int_0^{\infty} e^{\beta t}\pi e^{-\pi t}dt = \pi\int_0^{\infty} e^{(\beta-\pi)t}dt = \pi\left[\frac{e^{(\beta-\pi)t}}{\beta-\pi}\right]_0^{\infty} = \frac{\pi}{\beta-\pi}(0-1) = \frac{\pi}{\pi-\beta}, \beta<\pi.$$

d) $$\mathbb{E}[e^{\beta Z}|Y=n] = \mathbb{E}[e^{\beta S_n}|Y=n] = \mathbb{E}[e^{\beta S_n}] = \prod_{i=1}^{n}\mathbb{E}[e^{\beta X_i}] = \prod_{i=1}^{n}\frac{\pi}{\pi-\beta} = \left(\frac{\pi}{\pi-\beta}\right)^n, \beta<\pi$$

$$\Rightarrow \mathbb{E}[e^{\beta Z}|Y] = \left(\frac{\pi}{\pi-\beta}\right)^Y$$

$$\Rightarrow \mathbb{E}[e^{\beta Z}] = \mathbb{E}\left[\left(\frac{\pi}{\pi-\beta}\right)^Y\right] = \frac{\frac{\pi}{\pi-\beta}}{2-\frac{\pi}{\pi-\beta}} = \frac{\pi}{\pi-2\beta}, \quad \left|\frac{\pi}{\pi-\beta}\right|<2, \ \beta<\pi$$

Then $\left|\frac{\pi}{\pi-\beta}\right|<2 \Rightarrow \pi<2\pi-2\beta \Rightarrow 2\beta<\pi \Rightarrow \beta<\frac{\pi}{2}$

So, $\mathbb{E}[e^{\beta Z}] = \frac{\pi}{\pi-2\beta}, \ \beta<\frac{\pi}{2}$.

e) We are allowed to take derivative in expectation above. So, taking derivative wrt β three times,

$$\mathbb{E}[Z^3e^{\beta Z}] = \frac{d^3}{d\beta^3}\left(\frac{\pi}{\pi-2\beta}\right) = \frac{d^2}{d\beta^2}\left(\frac{2\pi}{(\pi-2\beta)^2}\right) = \frac{d}{d\beta}\left(\frac{8\pi}{(\pi-2\beta)^3}\right) = \frac{48\pi}{(\pi-2\beta)^4}$$

Then, setting $\beta=0$, we get

$$\mathbb{E}[Z^3] = \frac{48}{\pi^3}$$

* Note that from here we find $\mathbb{E}[Ze^{\beta Z}] = \frac{2\pi}{(\pi-2\beta)^2} \Rightarrow \mathbb{E}[Z] = \frac{2}{\pi}$ for part (a)

MATH 505a **QUALIFYING EXAM** **September 23, 2008**

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1.) The number of cars arriving at a McDonald's drive-up window in a given day is a Poisson random variable, N, with parameter λ. The numbers of passengers in these cars are independent random variables, X_i, each equally likely to be one, two, three, or four.

a) Simplify the probability generating function of N, say $G_N(z) = \mathbb{E}\, z^N =$ _____.

b) Simplify the moment generating function of $X = X_i$, say $M_X(t) = \mathbb{E}\, e^{tX} =$ _____.

c) Find the moment generating function of the total number of passengers passing by the drive-up window in a given day. (Hint: $S = \sum_{i=1}^{N} X_i$.)

2.) Consider a lottery with n^2 tickets, of which exactly n win prizes. A person buys $2n$ tickets. Find the following limits; part credit for guessing plausibly, and part credit for a proof.[Hint: the lottery involves drawing without replacement, but a good guess arises by thinking of drawing tickets with replacement.]

a) $\lim_{n\to\infty} \mathbb{P}($ at least one winning ticket$) =$ _____

b) $\lim_{n\to\infty} \mathbb{P}($ exactly 3 winning tickets$) =$ _____

3a) Suppose that S and S' are iid standard exponential, and $r > 0$.
Show that $P(rS < S') = 1/(1+r)$. [Hint: you might answer either by a detailed calculation, or by an informal Poisson process argument.]

3b) Find the covariance of (D_1, D_2, D_3, D_4) where $D_1 = (X_2 - X_1)/2, D_2 = (Y_2 - Y_1)/2, D_3 = X_3 - (X_1 + X_2)/2, D_4 = Y_3 - (Y_1 + Y_2)/2$ and the six coordinates (X_i, Y_i) for $i = 1$ to 3 are iid standard normal.

3c). Suppose that A, B, C are points in the plane, whose six coordinates (X_i, Y_i) for $i = 1$ to 3 are iid standard normal. There is a unique circle having the segment from A to B as a diameter. Show that the probability that C lies inside this circle is $1/4$. [Hint: 3a) and 3b) are both useful here. Think about the square of the distance from C to the midpoint of A,B, and about the square of the radius of the circle.]

Fall 2008, Applied Probability

1. a) $G_N(z) = \mathbb{E}[z^N] = \sum_{n=0}^{\infty} z^n e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda z)^n}{n!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}$.

b) $M_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=1}^{4} e^{tk} \frac{1}{4} = \frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4}$

c) $\mathbb{E}[e^{tS}] = \mathbb{E}[\mathbb{E}[e^{tS} | N]]$

$$\mathbb{E}[e^{tS} | N=n] = \mathbb{E}[e^{t(X_1 + \dots + X_n)} | N=n] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}] = \left(\frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4}\right)^n$$

So, $\mathbb{E}[e^{tS} | N] = \left(\frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4}\right)^N$ and thus,

$$M_S(t) = \mathbb{E}[e^{tS}] = \mathbb{E}\left[\left(\frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4}\right)^N\right] = e^{\lambda\left(\frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4} - 1\right)}$$

2. Since we are going to take limits, we can think that tickets are drawn with replacement. Because for large n the two cases give very similar results in means of probabilities.

a) P(at least one winning ticket) = 1 − P(no winning ticket)

$$= 1 - \left(\frac{n^2 - n}{n^2}\right)^{2n}$$

$$= 1 - \left(1 - \frac{1}{n}\right)^{2n}$$

$\lim_{n\to\infty}$ P(at least one winning ticket) $= 1 - \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^{2n} = 1 - e^{-2}$

b) Now, we can think that (since we assume "with replacement"), number of chosen winning tickets is a binomial r.v. There are $2n$ trials and the success probability is $\frac{n}{n^2} = \frac{1}{n}$. So,

$$P(\text{exactly 3 winning tickets}) = \binom{2n}{3} \left(\frac{1}{n}\right)^3 \left(\frac{n-1}{n}\right)^{2n-3}$$

$$= \underbrace{\frac{2n(2n-1)(2n-2)}{6} \cdot \frac{1}{n^3}}_{\to \frac{4}{3}} \cdot \underbrace{\left(1 - \frac{1}{n}\right)^{2n}}_{\to e^{-2}} \underbrace{\left(\frac{n}{n-1}\right)^3}_{\to 1} \quad \text{as } n \to \infty$$

So, $\lim_{n\to\infty}$ P(exactly 3 winning tickets) $= \frac{4e^{-2}}{3}$.

3. a) $P(rs < s') = \int_0^\infty \int_{rs}^\infty e^{-s} e^{-s'} \, ds' \, ds = \int_0^\infty e^{-s} \left[-e^{-s'} \right]_{rs}^\infty ds = \int_0^\infty e^{-s(r+1)} ds$

$= \left. \frac{e^{-s(r+1)}}{-(r+1)} \right|_0^\infty = -\frac{1}{r+1}(0-1) = \frac{1}{r+1}$

b) $D_1 = \frac{X_2 - X_1}{2}$, $D_2 = \frac{Y_2 - Y_1}{2}$, $D_3 = X_3 - \frac{X_1 + X_2}{2}$, $D_4 = Y_3 - \frac{Y_1 + Y_2}{2}$

$\rightarrow$ $Var(D_1) = \frac{1}{4}(Var X_2 + Var X_1) = \frac{1}{2}$ and $Var(D_2) = \frac{1}{2}$ (id)

$\rightarrow$ $Var(D_3) = 1 + \frac{1}{4}(1+1) = \frac{3}{2}$ and $Var(D_4) = \frac{3}{2}$ (id)

$\rightarrow$ $Cov(D_1, D_2) = Cov(D_1, D_4) = Cov(D_2, D_3) = Cov(D_3, D_4) = 0$ (independence)

$\rightarrow$ $Cov(D_1, D_3) = \frac{1}{2}\left[Cov\left(X_2, X_3 - \frac{X_1+X_2}{2}\right) - Cov\left(X_1, X_3 - \frac{X_1+X_2}{2}\right) \right]$

$= \frac{1}{2}\left[-\frac{1}{2} Cov(X_2, X_2) + \frac{1}{2} Cov(X_1, X_1) \right]$

$= 0$

$Cov(D_2, D_4) = 0$ (id, symmetry)

So, covariance of (D_1, D_2, D_3, D_4) is

$$\Sigma = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 3/2 & 0 \\ 0 & 0 & 0 & 3/2 \end{bmatrix}$$

c) We have $A(X_1, Y_1)$, $B(X_2, Y_2)$ and $C(X_3, Y_3)$. The center of the circle is $\left(\frac{X_1+X_2}{2}, \frac{Y_1+Y_2}{2} \right)$ and its radius is $\sqrt{\left(X_2 - \frac{X_1+X_2}{2}\right)^2 + \left(Y_2 - \frac{Y_1+Y_2}{2}\right)^2}$ $= \sqrt{D_1^2 + D_2^2}$. So, we want to find probability that the square of the distance between C and center of the circle is less than $D_1^2 + D_2^2$.

The distance from C to center is $\sqrt{\left(X_3 - \frac{X_1+X_2}{2}\right)^2 + \left(Y_3 - \frac{Y_1+Y_2}{2}\right)^2} = \sqrt{D_3^2 + D_4^2}$

So, we want to find $P(D_3^2 + D_4^2 < D_1^2 + D_2^2)$.

Because of the matrix that we found in part (b), we know that

D_1, D_2, D_3, D_4 are independent normal random variables. Remember that if we have uncorrelated multivariate normal random variables then they are independent. So, there exists U_1, U_2, U_3, U_4 iid $N(0,1)$ r.v. so that $D_1 = \frac{1}{\sqrt{2}} U_1$, $D_2 = \frac{1}{\sqrt{2}} U_2$, $D_3 = \sqrt{\frac{3}{2}} U_3$ and $D_4 = \sqrt{\frac{3}{2}} U_4$. So,

$$P(D_3^2 + D_4^2 < D_1^2 + D_2^2) = P\left(\frac{3}{2} U_3^2 + \frac{3}{2} U_4^2 < \frac{1}{2} U_1^2 + \frac{1}{2} U_2^2\right)$$

$$= P(3(U_3^2 + U_4^2) < U_1^2 + U_2^2)$$

Now, we know $U_i^2 \sim \chi_1^2$ for $i = 1, \ldots, 4$. Since they are all independent.

$U_3^2 + U_4^2 \sim \chi_2^2 = \text{Exponential}(2)$ and $U_1^2 + U_2^2 \sim \chi_2^2 = \text{Exponential}(2)$.

So, letting $U_3^2 + U_4^2 =: S_1$ and $U_1^2 + U_2^2 =: S_2$, we have

$$P(D_3^2 + D_4^2 < D_1^2 + D_2^2) = P(3S_1 < S_2)$$

where S_1 and S_2 are iid Exponential (2)

Now, again for $r > 0$ and $M, N \sim \text{Exponential}(\lambda)$,

$$P(rM < N) = \int_0^\infty \int_{rm}^\infty \frac{1}{\lambda^2} e^{-\frac{m}{\lambda}} e^{-\frac{n}{\lambda}} \, dn\, dm = \int_0^\infty \frac{1}{\lambda} e^{-\frac{m}{\lambda}} \left[-e^{-\frac{n}{\lambda}}\right]_{rm}^\infty dm$$

$$= \int_0^\infty \frac{1}{\lambda} e^{-\frac{m}{\lambda}(r+1)} dm = \left[\frac{e^{-\frac{m}{\lambda}(r+1)}}{-(r+1)}\right]_0^\infty = \frac{1}{r+1}$$

Thus, since S_1, S_2 are iid $\sim$ Exponential (2), we have (with $r = 3$)

$$P(D_3^2 + D_4^2 < D_1^2 + D_2^2) = P(3S_1 < S_2) = \frac{1}{4}$$

MATH 505a GRADUATE EXAM

SPRING 2008

Answer as many questions as you can. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. If you cannot do part (a) of a problem, you can still get credit for (b), (c) etc. by assuming the answer to (a). Start each problem on a fresh sheet of paper, and write on only one side of the paper.

(1) Let $Y \geq 0$ be a random variable with density f, and let X be another random variable. Assume X and Y have finite variances. Show that $E(XY) = \int_0^\infty E(XI_{[Y \geq t]})\, dt$. HINT: First express $E(XI_{[Y \geq t]})$ as an integral involving $E(X \mid Y = y)$.

(2) Let $\mathbf{X} = (X_1, .., X_m)$ and $\mathbf{Y} = (Y_1, .., Y_m)$ be random vectors with covariance matrices $\Sigma_{\mathbf{X}}$ and $\Sigma_{\mathbf{Y}}$.

(a) The *cross-covariance matrix* of $\mathbf{X}$ and $\mathbf{Y}$ is given by $C_{ij} = \operatorname{cov}(X_i, Y_j)$. For vectors $\mathbf{a}, \mathbf{b}$, express $\operatorname{var}(\mathbf{a} \cdot \mathbf{X} + \mathbf{b} \cdot \mathbf{Y})$ in terms of $\mathbf{a}, \mathbf{b}, \Sigma_{\mathbf{X}}, \Sigma_{\mathbf{Y}}$ and C.

(b) Suppose that for some vectors $\mathbf{a}, \mathbf{b}$ and some $k \in \mathbb{R}$, we have $\operatorname{var}((\mathbf{a}+u\mathbf{b}) \cdot \mathbf{X}) = ku$ for all $u \in \mathbb{R}$. Show that there are constants c_1, c_2 such that $P(\mathbf{a} \cdot \mathbf{X} = c_1) = P(\mathbf{b} \cdot \mathbf{X} = c_2) = 1$, and determine the value of k.

(3) Let U be a standard Cauchy random variable, that is, the density of U is $f_U(x) = \frac{1}{\pi}\frac{1}{1+x^2}, x \in \mathbb{R}$.

(a) Show that U and $1/U$ have the same distribution.

(b) Show that $E|U|^\alpha \geq 1$ for all $0 < \alpha < 1$. HINT: $1 = U \cdot \frac{1}{U}$.

(4) A sequence $X_1 X_2 \ldots X_n$ is said to have a local maximum at 1 if $X_1 > X_2$, a local maximum at i (for $1 < i < n$) if both $X_i > X_{i-1}$ and $X_i > X_{i+1}$, and a local maximum at n if $X_n > X_{n-1}$. Let N be the number of local maxima.

(a) Find the mean and variance for N in each of the following cases:

(i) if $X_1 X_2 \ldots X_n$ is a random permutation of the numbers $1, 2, \ldots, n$, with all $n!$ permutations equally likely;

Var(N) actually only cov(...) (ii) if $X_1, X_2, \ldots, X_n$ are chosen independently and uniformly from the integers $\{1, 2, \ldots, q\}$.

(b) Pick one of the two cases (i) or (ii) in (a), and show that $N/n \to 1/3$ in probability as $n \to \infty$.

Spring 2008, Applied Probability

1. $$\mathbb{E}[X 1_{\{Y\geq t\}}] = \int_0^\infty \mathbb{E}[X 1_{\{Y\geq t\}} \mid Y=y]\, f_Y(y)\,dy$$

$$= \int_0^\infty 1_{\{Y\geq t\}}\, \mathbb{E}[X \mid Y=y]\, f_Y(y)\,dy$$

$$= \int_t^\infty \mathbb{E}[X \mid Y=y]\, f_Y(y)\,dy$$

Then,

$$\int_0^\infty \mathbb{E}[X 1_{\{Y\geq t\}}]\,dt = \int_0^\infty \int_t^\infty \mathbb{E}[X\mid Y=y]\, f_Y(y)\,dy\,dt$$

since X and Y are square-integrable

$$= \int_0^\infty \int_0^y \mathbb{E}[X\mid Y=y]\, f_Y(y)\,dt\,dy$$

$$= \int_0^\infty y\, \mathbb{E}[X\mid Y=y]\, f_Y(y)\,dy$$

$$= \int_0^\infty \mathbb{E}[XY\mid Y=y]\, f_Y(y)\,dy$$

$$= \mathbb{E}\big[\mathbb{E}[XY\mid Y]\big]$$

$$= \mathbb{E}[XY]$$

2. a) $\operatorname{var}(a\cdot X + b\cdot Y) = \operatorname{var}(aX) + \operatorname{Var}(bY) + \operatorname{Cov}(aX, bY) + \operatorname{Cov}(bY, aX)$

$\rightarrow \operatorname{Cov}(aX, bY) = \sum_{i=1}^m \sum_{j=1}^m a_i b_j \operatorname{Cov}(X_i, Y_j) = \sum_{i=1}^m a_i * \left(\sum_{j=1}^m b_j C_{ij}\right) = a^T C b$

$\rightarrow \operatorname{Cov}(bY, aX) = \sum_{i=1}^m \sum_{j=1}^m b_i a_j \operatorname{Cov}(Y_i, X_j) = \sum_{i=1}^m b_i \left(\sum_{j=1}^m a_j C_{ji}\right) = b^T C^T a$

So, $\operatorname{Var}(a\cdot X + b\cdot Y) = a^T \Sigma_X a + b^T \Sigma_Y b + a^T C b + b^T C^T a$.

b) $$ku = \operatorname{var}((a+ub)\cdot X)$$

$$= (a+ub)^T \Sigma_X (a+ub)$$

$$= a^T \Sigma_X a + u\, a^T \Sigma_X b + u\, b^T \Sigma_X a + u^2 b^T \Sigma_X b$$

So, we obtain

$$a^T \Sigma_X a + (a^T \Sigma_X b + b^T \Sigma_X a - k) u + b^T \Sigma_X b u^2 = 0$$

Since this is valid $\forall u \in \mathbb{R}$, we must have

i) $a^T \Sigma_X a = 0 \Rightarrow var(a \cdot X) = 0 \Rightarrow \exists c_1 \in \mathbb{R}$ s.t. $P(a \cdot X = c_1) = 1$.

ii) $b^T \Sigma_X b = 0 \Rightarrow var(b \cdot X) = 0 \Rightarrow \exists c_2 \in \mathbb{R}$ s.t. $P(b \cdot X = c_2) = 1$

iii) $a^T \Sigma_X b + b^T \Sigma_X a - k = 0$.

Since $b^T \Sigma_X a \in \mathbb{R}$, we have $b^T \Sigma_X a = (b^T \Sigma_X a)^T = a^T \Sigma_X b$

So, we have $k = 2a^T \Sigma_X b$

3. a) For any nonzero $v \in \mathbb{R}$,

$$F_{\frac{1}{U}}(v) = P\left(\tfrac{1}{U} \le v\right) = P\left(U \ge \tfrac{1}{v}\right) = 1 - P\left(U < \tfrac{1}{v}\right) = 1 - F_U\left(\tfrac{1}{v}\right)$$

Then taking derivatives of both sides wrt v,

$$f_{\frac{1}{U}}(v) = -f_U\left(\tfrac{1}{v}\right) * \frac{-1}{v^2} = \frac{1}{v^2} \cdot \frac{1}{\pi} \cdot \frac{1}{1 + \frac{1}{v^2}} = \frac{1}{\pi} \cdot \frac{1}{1+v^2}$$

So, pdf of $1/U$ is the same as the pdf of U a.e. Thus, they have the same distribution.

b) Now consider by using Cauchy-Schwarz Inequality,

$$1 = \mathbb{E}\left[|U|^{\alpha/2} \frac{1}{|U|^{\alpha/2}}\right]^2 \le \mathbb{E}[|U|^\alpha] * \mathbb{E}\left[\left|\tfrac{1}{U}\right|^\alpha\right]$$

Since U and $\frac{1}{U}$ have the same distribution, we have $\mathbb{E}[|U|^\alpha] = \mathbb{E}\left[\left|\frac{1}{U}\right|^\alpha\right]$

So, we get $\mathbb{E}[|U|^\alpha]^2 \ge 1$ which implies $\mathbb{E}[|U|^\alpha] \ge 1$ since $|U|^\alpha > 0$

4 Let $N = 1_{\{X_1 > X_2\}} + 1_{\{X_n > X_{n-1}\}} + \sum_{i=2}^{n-1} 1_{\{X_i > X_{i-1}, X_i > X_{i+1}\}}$.

a) $\mathbb{E}[N] = P\{X_1 > X_2\} + P\{X_n > X_{n-1}\} + \sum_{i=2}^{n-1} P\{X_i > X_{i-1}, X_i > X_{i+1}\}$

$$= \frac{1}{2} + \frac{1}{2} + \sum_{i=2}^{n-1} \frac{1}{3}$$

$$= 1 + \frac{n-2}{3}$$

$$= \frac{n+1}{3}$$

$$\mathrm{Var}(N) = \mathrm{Var}(1_{\{X_1>X_2\}}) + \mathrm{Var}(1_{\{X_n>X_{n-1}\}}) + \sum_{i=2}^{n-1} \mathrm{Var}(1_{\{X_i>X_{i-1},\, X_i>X_{i+1}\}})$$

$$+2\,\mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_n>X_{n-1}\}}) + 2\sum_{i=2}^{n-1} \mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_i>X_{i-1},\, X_i>X_{i+1}\}})$$

$$+2\sum_{i=2}^{n-1} \mathrm{Cov}(1_{\{X_n>X_{n-1}\}}, 1_{\{X_i>X_{i-1},\, X_i>X_{i+1}\}})$$

$\rightarrow$ $\mathrm{Var}(1_{\{X_1>X_2\}}) = P\{X_1>X_2\} - P\{X_1>X_2\}^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

$\rightarrow$ $\mathrm{Var}(1_{\{X_n>X_{n-1}\}}) = \frac{1}{4}$

$\rightarrow$ $\mathrm{Var}(1_{\{X_i>X_{i-1},\, X_i>X_{i+1}\}}) = \frac{1}{3} - \frac{1}{9} = \frac{2}{9}$

$\rightarrow$ $\mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_n>X_{n-1}\}}) = P\{X_1>X_2,\, X_n>X_{n-1}\} - P\{X_1>X_2\}P\{X_n>X_{n-1}\} = \frac{1}{4} - \frac{1}{4} = 0$

$\rightarrow$ <u>If $i=2$:</u> $\mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_2>X_1,\, X_2>X_3\}}) = 0 - \frac{1}{2}\cdot\frac{1}{3} = -\frac{1}{6}$

$\rightarrow$ <u>If $i=3$:</u> $\mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_3>X_2,\, X_3>X_4\}}) = P\{X_1>X_2,\, X_3>X_2,\, X_3>X_4\} - P\{X_1>X_2\} * P\{X_3>X_2,\, X_3>X_4\}$

$* P\{X_1>X_2, X_3>X_2, X_3>X_4\} = P\{X_1>X_2,\, X_3>X_2>X_4\} + P\{X_1>X_2,\, X_3>X_4>X_2\}$

$= P\{X_1>X_3>X_2>X_4\} + P\{X_3>X_1>X_2>X_4\} + P\{X_1>X_3>X_4>X_2\}$

$+ P\{X_3>X_1>X_4>X_2\} + P\{X_3>X_4>X_1>X_2\}$

$= \frac{5}{4!} = \frac{5}{24}$ since all orderings are equally likely.

So, $\mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_3>X_2,\, X_3>X_4\}}) = \frac{5}{24} - \frac{1}{2}\cdot\frac{2}{9} = \frac{7}{72}$

$\rightarrow$ <u>If $i\geq 4$:</u> $\mathrm{Cov}(1_{\{X_1>X_2\}}, 1_{\{X_i>X_{i-1},\, X_i>X_{i+1}\}}) = P\{X_1>X_2,\, X_i>X_{i-1},\, X_i>X_{i+1}\} - P\{X_1>X_2\} * P\{X_i>X_{i-1}, X_i>X_{i+1}\}$

$* P\{X_1>X_2, X_i>X_{i-1}, X_i>X_{i+1}\} = P\{X_1>X_2,\, X_i>X_{i-1}>X_{i+1}\} + P\{X_1>X_2,\, X_i>X_{i+1}>X_{i-1}\}$

$= P\{X_1>X_2 \mid X_i>X_{i-1}>X_{i+1}\} P\{X_i>X_{i-1}>X_{i+1}\}$

$+ P\{X_1>X_2 \mid X_i>X_{i+1}>X_{i-1}\} P\{X_i>X_{i+1}>X_{i-1}\}$

$= \frac{1}{2}\cdot\frac{2}{9} + \frac{1}{2}\cdot\frac{2}{9} = \frac{2}{9}$

So, Cov(1

$\rightarrow \operatorname{Cov}\left(1_{\{X_n > X_{n-1}\}}, 1_{\{X_i > X_{i-1}, X_i > X_{i+1}\}}\right) = \frac{1}{9}$, by symmetry

Thus,

$$\operatorname{Var}(N) = \frac{1}{4} + \frac{1}{4} + (n-2)\,\frac{2}{9} + 2\left(-\frac{1}{6} + \frac{7}{72} + (n-4)\cdot\frac{1}{9}\right) + 2\left(-\frac{1}{6} + \frac{7}{72} + (n-4)\cdot\frac{1}{9}\right)$$

$$= \frac{1}{2} + \frac{2(n-2)}{9} + 4\left(-\frac{5}{72} + \frac{n-4}{9}\right)$$

$$= \frac{1}{2} + \frac{2(n-2)}{9} - \frac{5}{18} + \frac{4(n-4)}{9}$$

$$= \frac{2}{9} + \frac{2n-4+4n-16}{9}$$

$$= \frac{6n-18}{9}$$

$$= \frac{2n}{3} - 2$$

b) $E[N] = P\{X_1 > X_2\} + P\{X_n > X_{n-1}\} + \sum_{i=2}^{n-1} P\{X_i > X_{i-1}, X_i > X_{i+1}\}$

$$= \frac{\binom{q}{2}}{q^2} + \frac{\binom{q}{2}}{q^2} + \sum_{i=2}^{n-1}\left[P\{X_i > X_{i-1} > X_{i+1}\} + P\{X_i > X_{i+1} > X_{i-1}\}\right]$$

$$= \frac{q(q-1)}{2q^2} * 2 + \sum_{i=2}^{n-1} 2P\{X_i > X_{i-1} > X_{i+1}\}$$

$$= \frac{q-1}{q} + 2(n-2)\cdot\frac{\binom{q}{3}}{q^3}$$

$$= \frac{q-1}{q} + 2(n-2)\,\frac{q(q-1)(q-2)}{6q^3}$$

$$= \frac{q-1}{q} + (n-2)\,\frac{(q-1)(q-2)}{3q^2}$$

$$\operatorname{Var}(N) = \operatorname{Var}\left(1_{\{X_1>X_2\}}\right) + \operatorname{Var}\left(1_{\{X_n > X_{n-1}\}}\right) + \sum_{i=2}^{n-1}\operatorname{Var}\left(1_{\{X_i > X_{i-1}, X_i > X_{i+1}\}}\right)$$

$$+ 2\operatorname{Cov}\left(1_{\{X_1>X_2\}}, 1_{\{X_n>X_{n-1}\}}\right) + 2\sum_{i=2}^{n-1}\operatorname{Cov}\left(1_{\{X_1>X_2\}}, 1_{\{X_i>X_{i-1}, X_i>X_{i+1}\}}\right)$$

$$+ 2\sum_{i=2}^{n-1}\operatorname{Cov}\left(1_{\{X_n>X_{n-1}\}}, 1_{\{X_i>X_{i-1}, X_i > X_{i+1}\}}\right).$$

$\rightarrow \operatorname{Var}\left(1_{\{X_1>X_2\}}\right) = P\{X_1>X_2\} - P\{X_1>X_2\}^2 = \frac{q-1}{2q}\left(1 - \frac{q-1}{2q}\right) = \frac{(q-1)(q+1)}{4q^2}$

$\rightarrow \operatorname{Var}\left(1_{\{X_n>X_{n-1}\}}\right) = \frac{(q-1)(q+1)}{\quad}$ by symmetry.

$\longrightarrow \mathrm{Var}\left(1_{\{X_i > X_{i-1},\, X_i > X_{i+1}\}}\right) = P\{X_i > X_{i-1}, X_i > X_{i+1}\} - P\{X_i > X_{i+1}, X_i > X_{i+1}\}^2$

$$= \frac{(q-1)(q-2)}{3q^2}\left(1 - \frac{(q-1)(q-2)}{3q^2}\right)$$

$$= \frac{(q-1)(q-2)}{3q^2} \; \frac{(2q-1)(q+2)}{3q^2}$$

$$= \frac{(2q-1)(q-1)(q-2)(q+2)}{9q^4}$$

$\longrightarrow \mathrm{Cov}\left(1_{\{X_1 > X_2\}}, 1_{\{X_n > X_{n-1}\}}\right) = 0$ since $X_1, \ldots, X_n$ are chosen independently.

$\longrightarrow$ If $i=2$: $\mathrm{Cov}\left(1_{\{X_1 > X_2\}}, 1_{\{X_2 > X_1,\, X_2 > X_3\}}\right) = 0 - \frac{q-1}{2q} \; \frac{(q-1)(q-2)}{3q^2}$

$$= -\frac{(q-1)^2(q-2)}{6q^3}$$

$\longrightarrow$ If $i=3$ $\mathrm{Cov}\left(1_{\{X_1 > X_2\}}, 1_{\{X_3 > X_2,\, X_3 > X_4\}}\right) = P\{X_1 > X_2, X_3 > X_2, X_3 > X_4\} - P\{X_1 > X_2\}\, P\{X_3 > X_2, X_3 > X_4\}$ *

* $P(X_1 > X_2, X_3 > X_2, X_3 > X_4) = P(X_1 > X_2, X_3 > X_2 > X_4) + P(X_1 > X_2, X_3 > X_4 > X_2)$
$+ P(X_1 > X_2, X_3 > X_2 = X_4)$

b) We know from (a), $\mathbb{E}[N] = \frac{n+1}{3}$. For any $\varepsilon > 0$,

$$P\left(\left|\frac{N}{n} - \frac{1}{3}\right| > \varepsilon\right) = P\left(\left|N - \frac{n}{3}\right| > n\varepsilon\right)$$

Chebychev's Inequality

$$\leq \frac{\mathbb{E}\left[\left(N - \frac{n}{3}\right)^2\right]}{n^2\varepsilon^2}$$

$$= \frac{\mathbb{E}\left[\left(N - \frac{n+1}{3} + \frac{1}{3}\right)^2\right]}{n^2\varepsilon^2}$$

$$= \frac{\mathbb{E}\left[\left(N - \frac{n+1}{3}\right)^2\right] + \frac{2}{3}\mathbb{E}\left[N - \frac{n+1}{3}\right] + \frac{1}{9}}{n^2\varepsilon^2}$$

$$= \frac{Var(N) + \frac{1}{9}}{n^2\varepsilon^2}$$

$$= \frac{2n}{3n^2\varepsilon^2} - \frac{2}{n^2\varepsilon^2} + \frac{1}{9n^2\varepsilon^2}$$

$$= \frac{2}{3n\varepsilon} - \frac{2}{n^2\varepsilon^2} + \frac{1}{9n^2\varepsilon^2}$$

Taking limit as $n \to \infty$, $P\left(\left|\frac{N}{n} - \frac{1}{3}\right| > \varepsilon\right) \to 0$ as $n \to \infty$. Thus $\frac{N}{n} \to \frac{1}{3}$ in probability as $n \to \infty$.

MATH 505a GRADUATE EXAM

3 last part

FALL 2007

Answer as many questions as you can. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. If you cannot do part (a) of a problem, you can still get credit for (b), (c) etc. by assuming the answer to (a). Start each problem on a fresh sheet of paper, and write on only one side of the paper.

(1) (Daniel Bernoulli, 1786) Of the $2n$ people in a collection of n couples, exactly m die, with all $\binom{2n}{m}$ possibilities equally likely. Find the expected number of surviving couples.

(2) Let X be a standard normal random variable, and let Y be independent of X with $P(Y = 1) = P(Y = -1) = 1/2$. Answer the following questions and *justify your answers*:

(a) Is the random variable $Z = XY$ normally distributed?

(b) Do X and Z have a nonzero correlation?

(c) Does (X, Z) have a bivariate normal distribution?

(3) A sequence of mean-0 random variables $(X_n)_{n\in\mathbb{N}}$ is called *weakly stationary* if there is a function ϕ such that

$$E[X_iX_j] = \phi(|j-i|) < \infty \qquad \text{for all } i, j,$$

in other words this expected value only depends on the difference $|j - i|$. Suppose that for some such sequence, we have $\phi(k) \to 0$ as $k \to \infty$. Show that the weak law of large numbers is valid, that is, $\frac{X_1+\cdots X_n}{n}$ converges to 0 in probability.

(4) Consider a branching process with immigration: each generation is supplemented by an "immigrant" with probability p. This means that the size Z_n of the nth generation satisfies

$$Z_{n+1} = I_{n+1} + \sum_{i=1}^{Z_n} X_i,$$

where $I_{n+1} = 1$ with probability p, 0 with probability $1-p$, and the family sizes X_i are i.i.d. with generating function $G(s)$. We assume Z_n, I_{n+1} and $\{X_i\}$ are independent. Let $G_n(s)$ be the generating function of Z_n and let $\mu_n = EZ_n$.

(a) Show that $G_{n+1}(s) = (ps + (1-p))G_n(G(s))$. HINT: Condition on Z_n.

(b) Show that $\mu_{n+1} = p + \mu_n\mu$.

(c) If $\{\mu_n\}$ converges to a finite limit μ_∞, then what is μ_∞, in terms of p and μ?

Fall 2007, Applied Probability

1. Let A_i be the event that i^{th} couple survives. Then define $N = \sum_{i=1}^{n} 1_{A_i}$ to be the number of surviving couples.

If $m > 2n-2$, then $\mathbb{E}[N] = 0$. So, suppose $m \leq 2n-2$. Then,

$$\mathbb{E}[N] = \sum_{i=1}^{n} P(A_i)$$

$$P(A_i) = \frac{\binom{2n-2}{m}}{\binom{2n}{m}} = \frac{(2n-2)(2n-3)\cdots(2n-2-m+1)}{\cancel{m!}} \quad \frac{\cancel{m!}}{2n(2n-1)\cdots(2n-m+1)}$$

$$= \frac{\cancel{(2n-2)}\cancel{(2n-3)}\cdots\cancel{(2n-m+1)}(2n-m)(2n-m-1)}{2n(2n-1)\cancel{(2n-2)}\cdots\cancel{(2n-m+1)}}$$

$$= \frac{(2n-m)(2n-m-1)}{2n(2n-1)}$$

So, $\mathbb{E}[N] = \sum_{i=1}^{n} \frac{(2n-m)(2n-m-1)}{2n(2n-1)} = \frac{(2n-m)(2n-m-1)}{2(2n-1)}$

2. a) $F_Z(z) = P(Z \leq z) = P(XY \leq z) = P(XY \leq z \mid Y=1)P(Y=1) + P(XY \leq z \mid Y=-1)P(Y=-1)$

$$= P(X \leq z \mid Y=1) * \frac{1}{2} + P(-X \leq z) * \frac{1}{2} = \frac{1}{2}\left(P(X \leq z) + P(X \geq -z)\right)$$

$$= \frac{1}{2}\left(P(X \leq z) + P(X \leq z)\right) = P(X \leq z) = F_X(z)$$

Since $F_Z(z) = F_X(z)$ $\forall z \in \mathbb{R}$, we deduce that Z is also normally distributed.

b) $\mathrm{Cov}(X, Z) = \mathbb{E}[XZ] - \mathbb{E}[X]\mathbb{E}[Z]$

$$= \mathbb{E}[X^2 Y] - \mathbb{E}[X]\mathbb{E}[XY]$$

$$= \mathbb{E}[X^2]\mathbb{E}[Y] - \mathbb{E}[X]^2\mathbb{E}[Y]$$

independence of X and Y

$$= \mathbb{E}[Y]\,\mathrm{Var}(X)$$

$$= \left(\frac{1}{2}(-1) + \frac{1}{2}(1)\right) * 1$$

$$= 0$$

Thus, X and Z are uncorrelated.

c) If X and Z have bivariate normal distribution, for any region or in particular rectangle $R \subset \mathbb{R}^2$, we must have $P((X,Z) \in R) > 0$. So, consider

$$P((X,Z) \in [0,1] \times [1,2]) = P(X \in [0,1], Z \in [1,2])$$

$$= P(X \in [0,1],\ XY \in [1,2])$$

$$= P(X \in [0,1],\ XY \in [1,2] \mid Y=1)P(Y=1) + P(X \in [0,1],\ XY \in [1,2] \mid Y=-1)P(Y=-1)$$

by independence of X and Y

$$= \underbrace{P(X \in [0,1],\ X \in [1,2])}_{=0} * \frac{1}{2} + \underbrace{P(X \in [0,1],\ -X \in [1,2])}_{=0} * \frac{1}{2}$$

$$= 0$$

Thus, X and Y does not have bivariate normal distribution

3 Firstly observe that $\mathbb{E}[X_1 + \dots + X_n] = 0 \quad \forall n$ Then, for $\varepsilon > 0$,

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - 0\right| \geq \varepsilon\right) = P(|X_1 + \dots + X_n| \geq n\varepsilon)$$

$$\leq \frac{\mathbb{E}[|X_1 + \dots + X_n|^2]}{n^2 \varepsilon^2}$$

Chebychev's Inequality.

$$= \frac{1}{n^2 \varepsilon^2} \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$\rightarrow \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} \operatorname{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{n} \phi(0) + 2 \sum_{1 \leq i < j \leq n} \phi(j-i)$$

$$= n\phi(0) + 2\left[(n-1)\phi(1) + (n-2)\phi(2) + \dots + 2\phi(n-2) + \phi(n-1)\right]$$

$$= n\phi(0) + 2\sum_{j=1}^{n-1} (n-j)\phi(j)$$

So, $P\left(\left|\frac{X_1 + \dots + X_n}{n} - 0\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\left(\frac{\phi(0)}{n} + \frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\phi(j)\right)$

4. a) $G_{n+1}(s) = \mathbb{E}[s^{Z_{n+1}}]$

$\rightarrow \mathbb{E}[s^{Z_{n+1}} \mid Z_n = n] = \mathbb{E}[s^{I_{n+1} + X_1 + X_2 + \dots + X_n} \mid Z_n = n]$ by independence

$= \mathbb{E}[s^{I_{n+1}}] * \prod_{i=1}^{n} \mathbb{E}[s^{X_i}]$

$= (sp + (1-p)) [G(s)]^n$

So, $\mathbb{E}[s^{Z_{n+1}} \mid Z_n] = (sp + (1-p)) [G(s)]^{Z_n}$ and then

$G_{n+1}(s) = \mathbb{E}[\mathbb{E}[s^{Z_{n+1}} \mid Z_n]]$

$= \mathbb{E}[(sp + (1-p)) [G(s)]^{Z_n}]$

$= (sp + (1-p)) \mathbb{E}[G(s)^{Z_n}]$

$= (sp + (1-p)) G_n(G(s))$

b) We know that for any r.v. X with probability generating function $G(s)$, $G'(1) = \mathbb{E}[X]$, and $G(1) = 1$. So,

$G_{n+1}'(s) = p\, G_n(G(s)) + (sp + (1-p))\, G_n'(G(s))\, G'(s)$

$\Rightarrow \mu_{n+1} = G_{n+1}'(1) = p\, G_n(1) + G_n'(1)\, G'(1) = p + \mu_n \cdot \mu$

Thus, $\mu_{n+1} = p + \mu_n \cdot \mu$

c) Suppose $\lim_{n\to\infty} \mu_n = \alpha$ then taking limit of both sides in $\mu_{n+1} = p + \mu_n \cdot \mu$,

we get $\alpha = p + \alpha\mu \Rightarrow \alpha(1-\mu) = p \Rightarrow \alpha = \frac{p}{1-\mu}$. Thus

$\lim_{n\to\infty} \mu_n = \frac{p}{1-\mu}$

MATH 505a QUALIFYING EXAM January 31, 2007

Answer all three questions. Partial credit will be awarded, but in the event that you can not fully solve a problem you should state clearly what it is you have done and what you have left out. Unacknowledged omissions, incorrect reasoning and guesswork will lower your score. Start each problem on a fresh sheet of paper, and write on only one side of the paper. If you find that a calculation leads to something impossible, such as a negative probability or variance, indicate that something is wrong, but show your work anyway.

1.) n balls are placed into d boxes at random, with all n^d possibilities equally likely. Assume $d > 8$. Let X be the number of empty boxes. Let D be the event that no box receives more than 1 ball. Let A be the event that boxes 1 and 2 are both empty, B be the event that boxes 3,4,5 are all empty, and C be the event that boxes 6,7,8 are all empty.

a) Calculate and simplify: $\mathbb{E}\,X =$ ________________

b) Calculate and simplify: $\mathrm{Var\,X} =$ ________________

c) $\mathbb{P}(A \cup B \cup C) =$ ________________

d) If both $n, d \to \infty$ together, what relation must they satisfy in order to have $\mathbb{P}(D) \to .1$?

2.) Suppose Z is Poisson with $\mathbb{E}\,Z = \lambda < 1$. Let $X = 2^Z, Y = Z!$. Compute and simplify each of the following:

a) $\mathbb{E}\,Z^2 =$ ________________
b) $\mathbb{E}\,Z^3 =$ ________________
c) $\mathbb{E}\,Y =$ ________________
d) $\mathbb{E}\,X =$ ________________
e) $\mathrm{Var\,X} =$ ________________

3) Suppose that X is a sum of indicator random variables, with $\mu = \mathbb{E}\,X = 10, \sigma^2 = \mathrm{Var\,X} = 7$. Let A be the event $\{X > 0\}$.

a) State Chebyshev's inequality, involving the variance and the distance to the mean.
b) Apply Chebyshev's inequality to get a lower bound on $\mathbb{P}(A)$.
c) State the Cauchy-Schwarz for $(\mathbb{E}\,(XY))^2$.
d) Apply Cauchy-Schwarz, with $Y = 1(X > 0)$, the indicator that X is strictly positive, to get a lower bound on $\mathbb{P}(A)$.

Spring 2007, Applied Probability

1. Same as #3 in Fall 2013

2. Let us find probability generating function of Z.

$$G_Z(s) = E[s^Z] = \sum_{k=0}^{\infty} s^k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(s\lambda)^k}{k!} = e^{-\lambda} \cdot e^{s\lambda} = e^{\lambda(s-1)}.$$

Then let us find moment generating function of Z

$$M_Z(t) = E[e^{tZ}] = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

a) $E[Z^2] = M_Z''(0)$

$$M_Z'(t) = e^{\lambda(e^t-1)} \lambda e^t = \lambda e^{\lambda(e^t-1)+t}$$

$$M_Z''(t) = \lambda e^{\lambda(e^t-1)+t} (\lambda e^t + 1) = \lambda^2 e^{\lambda(e^t-1)+2t} + \lambda e^{\lambda(e^t-1)+t}$$

So, $E[Z^2] = \lambda^2 + \lambda$

b) $E[Z^3] = M_Z'''(0)$

$$M_Z'''(t) = \lambda^2 e^{\lambda(e^t-1)+2t} (\lambda e^t + 2) + \lambda e^{\lambda(e^t-1)+t} (\lambda e^t + 1)$$

$$E[Z^3] = \lambda^2(\lambda+2) + \lambda(\lambda+1) = \lambda^3 + 2\lambda^2 + \lambda^2 + \lambda = \lambda^3 + 3\lambda^2 + \lambda$$

c) $E[Y] = E[Z!] = \sum_{k=0}^{\infty} k! e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} e^{-\lambda} \lambda^k = e^{-\lambda} \sum_{k=0}^{\infty} \lambda^k = e^{-\lambda} \cdot \frac{1}{1-\lambda} = \frac{e^{-\lambda}}{1-\lambda}$,

since $0 < \lambda < 1$.

d) $E[X] = E[2^Z] = G_Z(2) = e^{\lambda(2-1)} = e^{\lambda}$

e) $Var(X) = E[X^2] - E[X]^2$

$\rightarrow E[X^2] = E[(2^Z)^2] = E[4^Z] = G_Z(4) = e^{\lambda(4-1)} = e^{3\lambda}$

So, $Var(X) = e^{3\lambda} - e^{2\lambda}$.

3. We have $\mu = E[X] = 10$, $\sigma^2 = Var(X) = 7$ and $A = \{X > 0\}$

a) For any r.v. X,

$$P(|X - E[X]| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

b) By Chebychev's Inequality, we have for any $\varepsilon > 0$.

$$P(|X-10| \geq \varepsilon) \leq \frac{7}{\varepsilon^2} \Rightarrow P(|X-10| < \varepsilon) \geq 1 - \frac{7}{\varepsilon^2}$$

$$\Rightarrow P(X > 10-\varepsilon) \geq P(|X-10| < \varepsilon) \geq 1 - \frac{7}{\varepsilon^2}$$

Set $\varepsilon = 10$ to get

$$P(X>0) \leq 1 - \frac{7}{10^2} = 0.93$$

c) For the r.v. X and Y,

$$\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[Y^2]$$

d) We have $\mathbb{E}[X \mathbb{1}_{\{X>0\}}]^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[\mathbb{1}^2_{\{X>0\}}] = \mathbb{E}[X^2]\,P(A)$

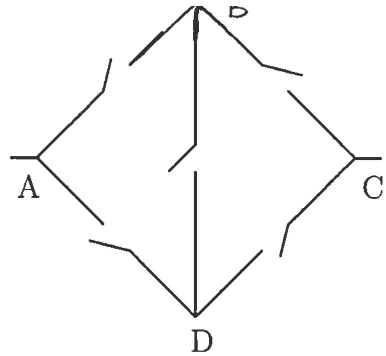
$\rightarrow$ $\mathbb{E}[X \mathbb{1}_{\{X>0\}}] = \mathbb{E}[X] - \underbrace{\mathbb{E}[X \mathbb{1}_{\{X=0\}}]}_{=0 \text{ since } X\mathbb{1}_{\{X=0\}}=0}$ since $X \geq 0$ by its definition.

So, $\mathbb{E}[X \mathbb{1}_{\{X>0\}}] = 10$

$\rightarrow$ $\mathbb{E}[X^2] = Var(X) + (\mathbb{E}[X])^2 = 7 + 100 = 107$

Substitution into the inequality yields

$$107\,P(A) \geq 100 \quad \Rightarrow \quad P(A) \geq \frac{100}{107}$$



3

FIGURE 1. A random connection

Fall 2006 Qualifying exam, Math 505a

Do all three problems, attempt all parts

Problem 1. Let X, Y, and Z be independent standard normal random variables.

(a) Show that $X^2 + Y^2$ and $\frac{X}{\sqrt{X^2+Y^2}}$ are independent.

(b) Show that

$$\frac{X + YZ}{\sqrt{1 + Z^2}}$$

is standard normal (Hint: condition on Z).

Problem 2. On Figure 1, each of the five connections can be open or closed independently of other connections. The probability to have a specific connection closed is p.

(a) Find the probability that there is a path of closed connections from A to C.

(b) Find the conditional probability that the connection along the diagonal BD is closed given that there is a path of closed connections from A to C.

Problem 3. Let S_n a random walk on $\mathbb{Z}$, with $S_0 = 0$. Let $\tau_0 = \inf\{n > 0 \colon S_n = 0\}$, the hitting time of 0.

(a) Show that

$$1 = \sum_{m=0}^{n} P_0(S_{n-m} = 0)\, P_0(\tau_0 > m).$$

(Hint: Condition according to the last time, that the chain will visit 0, before time n.)

(b) Assume further that S_n is *simple* random walk, that is, steps are plus one or minus one with probability one-half each. Assume also that n is even. The first term in the sum, indexed by $m = 0$, is simply $P(S_n = 0)$. Give a simple expression a_n which is asymptotic to this, that is, such that the ratio $a_n/P(S_n = 0)$ is close to 1 for large even n.

c) Continuing (b), the last term in the sum, indexed by $m = n$, is simply $P(\tau_0 > n)$. Give a simple expression b_n which is asymptotic to this.

Fall 2006, Applied Probability

1. a) Let $U = X^2 + Y^2$ and $V = \frac{X^2}{X^2+Y^2}$. We see that $U > 0$ and $0 < V < 1$

Consider

$$J = \begin{vmatrix} 2X & 2Y \\ \frac{2XY^2}{(X^2+Y^2)^2} & -\frac{2X^2Y}{(X^2+Y^2)^2} \end{vmatrix} = -\frac{4X^3Y + 4XY^3}{(X^2+Y^2)^2} = \frac{-4XY(X^2+Y^2)}{(X^2+Y^2)^2}$$

$$= \frac{-4XY}{X^2+Y^2}$$

$\Rightarrow X^2 = UV$ and $Y^2 = (1-V)U \Rightarrow X = \pm\sqrt{UV}$ and $Y = \pm\sqrt{(1-V)U}$

Then since $f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$, $-\infty < x < \infty$ and $-\infty < y < \infty$), we have

$$f_{U,V}(u,v) = \frac{1}{2\pi} e^{-\frac{1}{2}u} \frac{u}{4u\sqrt{v(1-v)}} \quad , u > 0, \quad 0 < v < 1$$

$$= \frac{1}{8\pi} e^{-\frac{1}{2}u} \frac{1}{\sqrt{v(1-v)}} \quad , u > 0, \quad 0 < v < 1$$

From their joint density, we see that U and V are independent. Thus, U and $\sqrt{V}$ are also independent, meaning that X^2+Y^2 and $\frac{X}{\sqrt{X^2+Y^2}}$ are independent.

b) Consider the r.v. $\frac{X+YZ}{\sqrt{1+Z^2}} \,|\, Z$. Since X, Y and Z are iid $N(0,1)$, given Z, $\frac{X+YZ}{\sqrt{1+Z^2}}$ has normal distribution. Let us find its mean and variance

$$\mathbb{E}\left[\frac{X+YZ}{\sqrt{1+Z^2}} \,\Big|\, Z\right] = \frac{1}{\sqrt{1+Z^2}} \mathbb{E}[X|Z] + \frac{Z}{\sqrt{1+Z^2}} \mathbb{E}[Y|Z] = \frac{1}{\sqrt{1+Z^2}} \mathbb{E}[X] + \frac{Z}{\sqrt{1+Z^2}} \mathbb{E}[Y] = 0$$

$$\mathbb{E}\left[\left(\frac{X+YZ}{\sqrt{1+Z^2}}\right)^2 \Big| Z\right] = \frac{1}{1+Z^2}\left(\mathbb{E}[X^2|Z] + 2Z\,\mathbb{E}[XY|Z] + Z^2\,\mathbb{E}[Y^2|Z]\right)$$

$$= \frac{1}{1+Z^2}\left(\mathbb{E}[X^2] + 2Z\,\mathbb{E}[X]\,\mathbb{E}[Y] + Z^2\,\mathbb{E}[Y^2]\right)$$

$$= \frac{1}{1+Z^2}(1+Z^2)$$

$$= 1$$

$$\mathrm{Var}\left(\frac{X+YZ}{\sqrt{1+Z^2}}\,\Big|\, Z\right) = \mathbb{E}\left[\left(\frac{X+YZ}{\sqrt{1+Z^2}}\right)^2 \Big| Z\right] - \mathbb{E}\left[\frac{X+YZ}{\sqrt{1+Z^2}}\,\Big|\, Z\right]^2 = 1-0 = 1$$

Thus, $\dfrac{X+YZ}{\sqrt{1+Z^2}}\,\Big|\, Z \sim N(0,1)$. Letting $A := \dfrac{X+YZ}{\sqrt{1+Z^2}}$, consider

$$f_{A,Z}(a,z) = f_{A|Z}(a|z)\, f_Z(z)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-a^2/2} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < a < \infty, \quad -\infty < z < \infty$$

So, A and Z are independent. Thus, $\dfrac{X+YZ}{\sqrt{1+Z^2}} \sim N(0,1)$

2. Let X be the event that there is no path from A to C.

a) We want to find $P(X^c)$. Now, consider that

$$P(X) = P(X \mid BD\text{ open})\,P(BD\text{ open}) + P(X \mid BD\text{ closed})\,P(BD\text{ closed})$$

$$= (1-p)\,P(X \mid BD\text{ open}) + p\,P(X \mid BD\text{ closed})$$

$\rightarrow P(X \mid BD\text{ open}) = pq(1-p^2) + q^2(1-p^2) + qp(1-p^2) = (1-p^2)^2$

$\rightarrow P(X \mid BD\text{ closed}) = p^2q^2 + 2pq^3 + q^4 + 2pq^3 + p^2q^2 = 2pq^2(p+2q) + q^4$

So, $P(X) = (1-p)(1-p^2)^2 + 2p^2q^2(p+2q) + pq^4$.

$$= q^3(1+p)^2 + 2p^2q^2(p+2q) + pq^4$$

$$= q^3 + 2pq^3 + p^2q^3 + 2p^3q^2 + 4p^2q^3 + pq^4$$

$$= 2p^3q^2 + (1+2p+5p^2)q^3 + pq^4$$

where $q = 1-p$. Thus $P(X^c) = 1 - 2p^3q^2 - (1+2p+5p^2)q^3 - pq^4$

b) $$P(BD\text{ closed} \mid X^c) = \frac{P(BD\text{ closed}, X^c)}{P(X^c)} = \frac{P(BD\text{ closed}) - P(BD\text{ closed}, X)}{P(X^c)}$$

$$= \frac{p - 2p^2q^2(p+2q) - pq^4}{1 - 2p^3q^2 - (1+2p+5p^2)q^3 - pq^4}$$

36

MATH 505a QUALIFYING EXAM February 1, 2006

Last Name: ________________ **First Name:**________________

ID#: ________________ **Signature:**________________

Fall 2004, #2 → better explanation of problem

1. A *run* in a sequence of coin tosses is a maximal subsequence of consecutive tosses all having the same outcome; for example HHHTHHTTH has 5 runs. A biased coin, with $p = \mathbf{P}(\text{heads}) \in (0,1)$, is tossed n times. Write $q = 1 - p$. Let R_n be the number of runs in the first n tosses. Find exact formulas for

a) $\mu_n = \mathbf{E}R_n$ and

b) $\sigma_n^2 = \text{Var}(R_n)$.

HINT: pay careful attention to boundary effects-what happens at the start and end of the sequence of n tosses. Note that $\mu_1 = 1, \sigma_1^2 = 0$, and use this as a check on your answers. Note also that $\mathbf{P}(R_2 = 1) = p^2 + q^2$, $\mathbf{P}(R_2 = 2) = 2pq$, so $\mu(2) = 1 + 2pq$. ← μ_2

c) For the special case $p = 1/2$, the distribution of $R_n - 1$ is very well known distribution (e.g. Binomial, Poisson, Hypergeometric, Geometric, etc) NAME the distribution AND its parameter(s).

2. Assume the vector $\mathbf{X} = (X_1, \ldots, X_N)$ has a multivariate normal distribution $N(\mu, \mathbf{V})$, where μ is the vector of expected values and $\mathbf{V}$ is the covariance matrix. Let $c_1, \ldots, c_N$ be constants.

Find the distribution of $Y = \sum_{i=1}^{N} c_i X_i$.

3. Let $S_n = X_1 + \cdots + X_n, n \geqq 1$, be a random walk, where $EX_k = \mu$ and $Var(X_k) = \sigma^2$, $0 < \sigma^2 < \infty$.

+ (a) Find the covariance $Cov(S_n, S_m)$ and the correlation coefficient $\rho(S_n, S_m)$ of S_n and S_m, $m \neq n$.

(b) Assume $n > m$. Find $\lim_{n\to\infty} Cov(S_n, S_m)$ and $\lim_{n\to\infty} \rho(S_n, S_m)$. Does $\lim_{n\to\infty} Cov(S_n, S_m)$ depend on the distribution of the increments? Does $\lim_{n\to\infty} \rho(S_n, S_m)$ depend on the distribution of the increments?

Spring 2006, Applied Probability.

1. Define A_i as the event that i^{th} and $(i+1)^{st}$ flip gives different outcomes (i.e. either HT or TH). So, $R_n = 1 + \sum_{i=1}^{n-1} 1_{A_i}$. Then,

a) $E[R_n] = 1 + \sum_{i=1}^{n-1} P(A_i) = 1 + \sum_{i=1}^{n-1} (pq + qp) = 1 + 2(n-1)pq$

b) $Var(R_n) = \sum_{i=1}^{n-1} var(1_{A_i}) + \sum_{1 \le i < j \le n-1} Cov(1_{A_i}, 1_{A_j})$

$\rightarrow Var(1_{A_i}) = P(A_i) - P(A_i)^2 = 2pq(1-2pq)$

$\rightarrow$ If $j = i+1$: $Cov(1_{A_i}, 1_{A_{i+1}}) = P(A_i \cap A_{i+1}) - P(A_i)P(A_{i+1}) = (pqp + qpq) - 4p^2q^2$

$= pq(p+q-4pq)$

$\rightarrow$ If $j \geq i+2$. $Cov(1_{A_i}, 1_{A_j}) = P(A_i \cap A_j) - P(A_i)P(A_j) = 0$ because of independence of trials.

So, $Var(R_n) = 2(n-1)pq(1-2pq) + (n-2)pq(p+q-4pq)$

c) Suppose $p=q=1/2$. Consider $P_n := R_n - 1 = \sum_{i=1}^{n-1} 1_{A_i}$. We know that when $|j-i| \geq 2$, 1_{A_i} and 1_{A_j} are independent. When $|j-i| = 1$, wlog when $j = i+1$, 1_{A_i} and 1_{A_j} are two random variables both taking only two values and $Cov(1_{A_i}, 1_{A_j}) = \frac{1}{4}(1 - 4 \cdot \frac{1}{4}) = 0$. So, they are actually independent. That is, P_n is summation of independent Bernoulli r.v. which means P_n has Binomial distribution.

$E[P_n] = 1 + 2(n-1) * \frac{1}{4} - 1 = \frac{n-1}{2} \quad (= (n-1)p)$

$Var(P_n) = 2(n-1) * \frac{1}{4}(1 - \frac{1}{2}) + (n-2) * \frac{1}{4}(1 - 4 * \frac{1}{4}) = \frac{n-1}{4} \quad (= (n-1)pq)$

Thus $P_n \sim$ Binomial$(n-1, 1/2)$

2. Since X has a multivariate normal distribution, we know that,

$X_i = \mu_i + \sum_{j=1}^{N} a_{ij} Z_j$ for $i = 1, \ldots, N$ where $Z_1, \ldots, Z_N$ are iid $N(0,1)$. So,

$Y = \sum_{i=1}^{N} c_i \mu_i + + \sum_{i=1}^{N} \sum_{j=1}^{N} c_i a_{ij} Z_j$ which means Y has normal distribution.

$$E[Y] = \sum_{i=1}^{N} c_i \mu_i = c^T \mu$$

$$Var(Y) = \sum_{i=1}^{N} c_i^2 V_{ii} + \sum_{1 \le i < j \le N} c_i c_j V_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} c_i c_j V_{ij} = c^T V c$$

Thus, $Y \sim Normal(c^T \mu, c^T V c)$.

2. a) Suppose wlog that $n > m$.

$$Cov(S_n, S_m) = \sum_{i=1}^{n} \sum_{j=1}^{m} Cov(X_i, X_j) = \sum_{j=1}^{m} Var(X_j) = m\sigma^2$$

$$Var(S_n) = \sum_{i=1}^{n} Var(X_i) = n\sigma^2$$ and similarly, $Var(S_m) = m\sigma^2$

$$\rho(S_n, S_m) = \frac{Cov(S_n, S_m)}{\sqrt{Var(S_n)}\sqrt{Var(S_m)}} = \frac{m\sigma^2}{\sqrt{n}\sigma\sqrt{m}\sigma} = \frac{\sqrt{m}}{\sqrt{n}}$$

MATH 505a QUALIFYING EXAM September 13, 2005

Last Name: ____________ **First Name:** ____________

ID#: ____________ **Signature:** ____________

1. A building has $k > 1$ floors above the street level.

Some number $m > 1$ of people enter at street level and board an elevator. They choose floors independently, uniformly at random. Let X_i be the indicator of the event that at least one person selects floor i, so that $S = X_1 + \ldots + X_k$ is the number of stops the elevator must make.

a) Find the constant $a = \mathbf{E}X_1$ as a function of m, k.

b) Find the constant $b = \mathbf{E}X_1X_2$ as a function of m, k.

c) Find $c = \text{Cov}(X_1, X_2)$ in terms of a, b.

d) Use the method of indicators to find $\mathbf{E}S$,

e) Use the variance-covariance expansion to find $\text{Var}(S)$ in terms of a, b, m.

2. Let $(X_n, n \geq 0)$ and $(Y_n, n \geq 0)$ be two independent simple random walks on $\mathbb{Z}$ starting at zero.

a) Prove that the sequence of ordered pairs $(Z_n = (X'_n, Y'_n), n \geq 0)$, where

$$X'_n = \frac{X_n + Y_n}{2} \text{ and } Y'_n = \frac{X_n - Y_n}{2},$$

is a simple random walk on $\mathbb{Z}^2$.

b) For a standard basis vector $\mathbf{e}_i$ in $\mathbf{R}^2$, find $\mathbf{P}(Z_1 = -\mathbf{e}_i)$.

c) Find $\mathbf{P}(Z_n = 0)$.

d) Find the asymptotic of $\mathbf{P}(Z'_{2n} = 0)$ as $n \longrightarrow \infty$.

Hint: Use Stirling's formula: $n! \sim n^n e^{-n}\sqrt{2\pi n}$.

3. Let X have the binomial distribution $\text{Bin}(n, U)$, where U is uniform on $(0, 1)$. Show that X is uniformly distributed on $\{0, 1, \ldots, n\}$.

Hint. Compute the generating function of X.

4. The n randomly chosen real numbers $a_1, \ldots, a_n$ are rounded off to the nearest hundredth $a_1 + X_1, \ldots a_n + X_n$, where the round-off errors $X_1, \ldots, X_n$ are assumed to be independent and uniformly distributed on $[-\frac{1}{2}10^{-2}, \frac{1}{2}10^{-2}]$.

Use the CLT to find a number $\lambda > 0$ (depending upon n) such that

$$\mathbf{P}(\sum_{i=1}^{n} |X_i| < \lambda) \approx 0.99.$$

STANDARD NORMAL DISTRIBUTION FUNCTION

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,dt, \qquad x \geq 0.$$

For $x < 0$, use the relation $\Phi(x) = 1 - \Phi(-x)$.

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Fall 2005, Applied Probability

1. a) $a = E[X_1] = P(X_1 = 1) = 1 - P(X_1 = 0) = 1 - P(\text{no one choses } 1^{st} \text{ floor})$

$= 1 - \left(\frac{k-1}{k}\right)^m$

b) $b = E[X_1 X_2] = P(X_1 = 1, X_2 = 1)$

$\rightarrow P(X_1 = 1, X_2 = 1) = 1 - P(X_1 = 0, X_2 = 1) - P(X_1 = 1, X_2 = 0) - P(X_1 = 0, X_2 = 0)$

$* \; P(X_1 = 0, X_2 = 1) = P(X_1 = 0) - P(X_1 = 0, X_2 = 0)$

$= \left(\frac{k-1}{k}\right)^m - \left(\frac{k-2}{k}\right)^m$

Note that $P(X_1 = 1, X_2 = 0) = P(X_1 = 0, X_2 = 1)$ by symmetry.

$* \; P(X_1 = 0, X_2 = 0) = \left(\frac{k-2}{k}\right)^m$

So, $b = P(X_1 = 1, X_2 = 1) = 1 - 2\left(\frac{k-1}{k}\right)^m - 2\left(\frac{k-2}{k}\right)^m - \left(\frac{k-2}{k}\right)^m$

$= 1 - 2\left(\frac{k-1}{k}\right)^m - 3\left(\frac{k-2}{k}\right)^m$

c) $c = Cov(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2]$

$= 1 - 2\left(\frac{k-1}{k}\right)^m - 3\left(\frac{k-2}{k}\right)^m - \left[1 - \left(\frac{k-1}{k}\right)^m\right]^2$

$= -\left(\frac{k-1}{k}\right)^m \left[3 + \left(\frac{k-1}{k}\right)^m\right] \quad = b - a^2$

d) $E[S] = \sum_{i=1}^{k} E[X_i] = \sum_{i=1}^{k} P(X_i = 1) = k - k\,\frac{(k-1)^m}{k^m} = k - \frac{(k-1)^m}{k^{m-1}}$.

e) $Var(S) = \sum_{i=1}^{k} Var(X_i) + \sum_{1 \le i < j \le k} Cov(X_i, X_j)$

$= \sum_{i=1}^{k} (a - a^2) + \sum_{1 \le i < j \le k} (b - a^2)$

$= k(a - a^2) + \frac{k(k-1)}{2}(b - a^2)$

2. a) We see that

$X_n' = \frac{X_n + Y_n}{2} = \begin{cases} 1 & wp \; 1/4 \\ 0 & wp \; 1/2 \\ -1 & wp \; 1/4 \end{cases}$ and $Y_n' = \frac{X_n - Y_n}{2} = \begin{cases} 1 & wp \; 1/4 \\ 0 & wp \; 1/2 \\ -1 & wp \; 1/4 \end{cases}$

But they cannot take same of the values at the same time.

$$P(X_n'=1, Y_n'=1) = P(X_n+Y_n=2, X_n-Y_n=2) = 0$$

$$P(X_n'=1, Y_n'=0) = P(X_n+Y_n=2, X_n-Y_n=0) = 1/4$$

$$P(X_n'=1, Y_n'=-1) = P(X_n+Y_n=2, X_n-Y_n=-2) = 0$$

$$P(X_n'=0, Y_n'=1) = P(X_n+Y_n=0, X_n-Y_n=2) = 1/4$$

$$P(X_n'=0, Y_n'=0) = P(X_n+Y_n=0, X_n-Y_n=0) = 0$$

$$P(X_n'=0, Y_n'=-1) = P(X_n+Y_n=0, X_n-Y_n=-2) = 1/4$$

$$P(X_n'=-1, Y_n'=1) = P(X_n+Y_n=-2, X_n-Y_n=2) = 0$$

$$P(X_n'=-1, Y_n'=0) = P(X_n+Y_n=-2, X_n-Y_n=0) = 1/4$$

$$P(X_n'=-1, Y_n'=-1) = P(X_n+Y_n=-2, X_n-Y_n=-2) = 0$$

Thus, we have

$$Z_n = \begin{cases} (1,0) & wp \ 1/4 \\ (0,1) & wp \ 1/4 \\ (0,-1) & wp \ 1/4 \\ (-1,0) & wp \ 1/4 \end{cases}$$

and since $(X_n)_{n\geq 0}, (Y_n)_{n\geq 0}$ are independent, $(Z_n)_{\geq 0}$ is also independent. Thus $(Z_n)_{n\geq 0}$ is a simple random walk on $\mathbb{Z}^2$.

b) From above definition of Z_n, for any standard basis vector e_i in $\mathbb{R}^2$ (which are $(0,1),(1,0)$) we have $P(Z_1=e_i)=1/4$.

c) From above calculations, $P(Z_n=0) = P(X_n+Y_n=0, X_n-Y_n=0) = P(X_n=0, Y_n=0) = 0$.

d) Assuming $Z_n' = \sum_{k=1}^{n} Z_k$, consider

$$P(Z_{2n}'=0) = P\left(\sum_{k=1}^{2n} Z_k = 0\right) = P\left(\sum_{k=1}^{2n} X_k'=0, \sum_{k=1}^{2n} Y_k'=0\right)$$

$$= P\left(\sum_{k=1}^{2n} X_k + \sum_{k=1}^{2n} Y_k = 0, \sum_{k=1}^{2n} X_k - \sum_{k=1}^{2n} Y_k = 0\right) = P\left(\sum_{k=1}^{2n} X_k = 0, \sum_{k=1}^{2n} Y_k = 0\right)$$

$$\underset{\substack{\uparrow \\ \text{by independence}}}{=} P\left(\sum_{k=1}^{2n} X_k = 0\right) P\left(\sum_{k=1}^{2n} Y_k = 0\right) \underset{\substack{\uparrow \\ \text{identically}}}{=} P\left(\sum_{k=1}^{2n} X_k = 0\right)^2$$

4. Firstly let $U \sim \text{Uniform}[-a,a]$, let us find mean and variance of $|U|$

$$\mathbb{E}[|U|] = \int_{-a}^{a} |u| \frac{1}{2a} du = \frac{1}{2a}\left[\int_{-a}^{0} -u\,du + \int_{0}^{a} u\,du \right] = \frac{1}{2a} \times 2 \int_{0}^{a} u\,du = \frac{1}{a}\left[\frac{u^2}{2}\right]_0^a$$

$$= \frac{a}{2}$$

$$\mathbb{E}[|U|^2] = \mathbb{E}[U^2] = Var(U) + \mathbb{E}[U]^2 = \frac{(a+a)^2}{12} + 0 = \frac{a^2}{3}$$

So, $Var(|U|) = \frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12}$

So, we have $\mathbb{E}\left[\sum_{i=1}^{n} |X_i|\right] = \sum_{i=1}^{n} \frac{1}{4} 10^{-2} = \frac{n}{4} 10^{-2}$ and

$Var\left(\sum_{i=1}^{n} |X_i|\right) = \sum_{i=1}^{n} \frac{1}{48} 10^{-4} = \frac{n}{48} 10^{-4}$. Then by CLT, we know

$$\frac{\sum_{i=1}^{n} |X_i| - \frac{n}{4} 10^{-2}}{\sqrt{\frac{n}{48}}\, 10^{-2}} \xrightarrow{d} N(0,1)$$

So,

$$P\left(\sum_{i=1}^{n} |X_i| < \lambda\right) = P\left(\frac{\sum_{i=1}^{n} |X_i| - \frac{n}{4} 10^{-2}}{\sqrt{\frac{n}{48}}\, 10^{-2}} < \frac{\lambda - \frac{n}{4} 10^{-2}}{\sqrt{\frac{n}{48}}\, 10^{-2}}\right)$$

$$\approx P\left(Z < \frac{\lambda - \frac{n}{4} 10^{-2}}{\sqrt{\frac{n}{48}}\, 10^{-2}}\right)$$

where $Z \sim N(0,1)$. We want to find λ so that

$$P\left(Z < \frac{\lambda - \frac{n}{4} 10^{-2}}{\sqrt{\frac{n}{48}}\, 10^{-2}}\right) \approx 0.99$$

Since $Z \sim N(0,1)$, from table, we get $\frac{\lambda - \frac{n}{4} 10^{-2}}{\sqrt{\frac{n}{48}}\, 10^{-2}} \approx 2.325$ and so,

$$\lambda \approx 0.2325 \frac{\sqrt{n}}{4\sqrt{3}} + n\,0.0025$$

$$P\left(\sum_{k=1}^{2n} X_k = 0\right) = \binom{2n}{n}\frac{1}{2^n}\frac{1}{2^n} = \binom{2n}{n}\frac{1}{2^{2n}}$$

Then, by Stirling's formula,

$$P(Z'_{2n}=0) = \binom{2n}{n}^2 \frac{1}{2^{4n}} \approx \left[\frac{(2n)^{2n}\cancel{e^{-2n}}\sqrt{4\pi n}}{n^{2n}\cancel{e^{-2n}}\,2\pi n}\right]^2 \frac{1}{2^{4n}} = \frac{2^{4n}}{\pi n}\ \frac{1}{2^{4n}} = \frac{1}{\pi n}$$

Thus, $P(Z'_{2n}=0) \longrightarrow 0$ as $n \to \infty$

3. By definition $G_X(s) = \mathbb{E}[s^X]$. Firstly consider that

$$E[s^X \mid U=u] = \sum_{k=0}^{n} s^k P(X=k \mid U=u)$$

$$= \sum_{k=0}^{n} s^k \binom{n}{k} u^k (1-u)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (su)^k (1-u)^{n-k}$$

$$= (1+u(s-1))^n$$

So, $E[s^X|U] = (1+U(s-1))^n$ Then,

$$\mathbb{E}[\mathbb{E}[s^X|U]] = \mathbb{E}[(1+U(s-1))^n]$$

$$= \int_0^1 (1+u(s-1))^n du$$

$$= \left[\frac{(1+u(s-1))^{n+1}}{(n+1)(s-1)}\right]_0^1$$

$$= \frac{s^{n+1}-1}{(n+1)(s-1)}$$

So, we have $G_X(s) = \mathbb{E}[s^X] = \mathbb{E}[\mathbb{E}[s^X|U]] = \dfrac{1}{n+1}\ \dfrac{s^{n+1}-1}{s-1}$

On the other hand let Y be a uniformly distributed r.v. on $\{0,1,\dots,n\}$
The probability generating function of Y is

$$G_Y(s) = \mathbb{E}[s^Y] = \sum_{k=0}^{n} s^k P\{Y=k\} = \frac{1}{n+1}\sum_{k=0}^{n} s^k = \frac{1}{n+1}\ \frac{s^{n+1}-1}{s-1}$$

Since $G_X(s) = G_Y(s)$, X and Y have the same distribution. That is, X is uniformly distributed on $\{0,1,\dots,n\}$.

MATH 505a QUALIFYING EXAM February 03, 2005

Last Name: ________________ **First Name:**________________

ID#: ________________ **Signature:**________________

You should try at least three problems; you may try all four.

1. Let $X_1, X_2, \ldots, X_n$ be a sample of independent, identically distributed random variables with density f and distribution function F. Let $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ be the ordered sequence of those variables. For $i < j$ find the joint density of $(X_{(i)}, X_{(j)})$.

2. Let N be a Poisson random variable with parameter λ. Let $Y = \sum_{i=1}^{N} X_i$, where X_i are independent, identically distributed, non-negative integer valued random variables with finite mean. Show that for any function g (such that the expectations exist) we have

$$E[Yg(Y)] = \lambda E[X_0 g(Y + X_0)].$$

3. A stick is broken in two pieces, uniformly at random. Let X denote the ratio of the lengths of the shorter to the longer piece. Find the mean and the variance of X.

4. The number of the electrons that hit the plate is Poisson with parameter $\lambda_1 = 2$. Every impact produces independently a number of secondary electrons that is Poisson with parameter $\lambda_2 = 1$. a) Find the moment generating function of the total number of secondary electrons; b) Find the variance of that number.

Spring 2005, Applied Probability:

1. The joint density of $(X_{(i)}, X_{(j)})$ for $i<j$ is as follows

$$f_{X_{(i)},X_{(j)}}(u,v) = \frac{n!}{(i-1)!\,(n-j-1)!\,(n-j)!}[F(u)]^{i-1} f(u)\,[F(v)-F(u)]^{j-i-1} f(v)\,[1-F(v)]^{n-j}$$

2. Firstly consider that (assuming X_i's and N are all independent)

$$\mathbb{E}[Yg(Y)\,|\,N=n] = \mathbb{E}\Big[\sum_{i=1}^{n} X_i\, g\Big(\sum_{i=1}^{n} X_i\Big)\,\Big|\,N=n\Big] = \sum_{i=1}^{n}\mathbb{E}\Big[X_i\, g\Big(\sum_{j=1}^{n} X_j\Big)\Big]$$

Then,

$$\mathbb{E}[Yg(Y)] = \mathbb{E}[\mathbb{E}[Yg(Y)\,|\,N]]$$

$$= \sum_{k=1}^{\infty} e^{-\lambda}\cdot\frac{\lambda^k}{k!}\Big(\underbrace{\sum_{i=1}^{k}\mathbb{E}\Big[X_i\, g\Big(\sum_{j=1}^{k} X_j\Big)\Big]}\Big)$$

$$= \mathbb{E}\Big[X_0\, g\Big(\sum_{j=1}^{k-1} X_j + X_0\Big)\Big]$$ since X_i's are identically distributed

$(= \mathbb{E}[X_0 g(X_0)]$ if $k=1)$

$$= \sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\, k\, \mathbb{E}\Big[X_0\, g\Big(\sum_{j=1}^{k-1} X_j + X_0\Big)\Big]$$

$$= \lambda\sum_{k=1}^{\infty} e^{-\lambda}\cdot\frac{\lambda^{k-1}}{(k-1)!}\,\mathbb{E}\Big[X_0\, g\Big(\sum_{j=1}^{k-1} X_j + X_0\Big)\Big]$$

$$= \lambda\sum_{k=0}^{\infty} e^{-\lambda}\cdot\frac{\lambda^{k}}{k!}\,\mathbb{E}\Big[X_0\, g\Big(\sum_{j=1}^{k} X_j + X_0\Big)\Big]$$

$$= \lambda\,\mathbb{E}\big[\mathbb{E}[X_0\, g(Y+X_0)]\,|\,N\big]$$

$$= \lambda\,\mathbb{E}[X_0\, g(Y+X_0)]$$

3. Let us denote the distance between the midpoint of the stick and the point where it is broken by Y. Also, let ℓ be the length of the stick. Then,

$$X = \frac{\ell/2 - Y}{\ell/2 + Y} = -1 + \frac{\ell}{\ell/2 + Y}$$

Since the stick is broken uniformly, we see that $Y \sim \text{Uniform}[0, \ell/2]$.
Now, consider

$$\mathbb{E}[X] = -1 + \ell\, \mathbb{E}\left[\frac{1}{\ell/2 + Y}\right]$$

$$= -1 + \ell \int_0^{\ell/2} \frac{1}{\ell/2 + y} \cdot \frac{2}{\ell}\, dy$$

$$= -1 + 2\left[\log(\ell/2 + y)\right]_0^{\ell/2}$$

$$= -1 + 2\left(\log(\ell) - \log(\ell/2)\right)$$

$$= -1 + 2\log 2$$

$$\mathbb{E}[X^2] = 1 + 2\ell\, \mathbb{E}\left[\frac{1}{\ell/2 + Y}\right] + \ell^2\, \mathbb{E}\left[\frac{1}{(\ell/2 + Y)^2}\right]$$

$$= -1 + 4\log 2 + \ell^2 \int_0^{\ell/2} \frac{1}{(\ell/2 + y)^2} \frac{2}{\ell}\, dy$$

$$= -1 + 4\log 2 + 2\ell\left[-\frac{1}{\ell/2 + y}\right]_0^{\ell/2}$$

$$= -1 + 4\log 2 + 2$$

$$= 1 + 4\log 2$$

So, $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \cancel{1} + 4\log 2 - \cancel{1} + 4\log 2 + 4(\log 2)^2 = 4\log 2(2 + \log 2)$

4. Let N be the number of electrons that hit the plate where $N \sim \text{Poisson}(2)$ and let T_i be the number of secondary electrons produced by ith electron where $T_i \sim \text{Poisson}(1)$. Let T be the total number of secondary electrons. So, we have $T = \sum_{i=1}^{N} T_i$.

a) $\rightarrow \mathbb{E}[e^{tT_i}] = \sum_{k=0}^{\infty} e^{tk} e^{-1} \frac{1}{k!} = e^{-1} \sum_{k=0}^{\infty} \frac{(e^t)^k}{k!} = e^{e^t - 1}$

$\rightarrow \mathbb{E}[s^N] = \sum_{k=0}^{\infty} s^k e^{-2} \frac{2^k}{k!} = e^{-2} \sum_{k=0}^{\infty} \frac{(2s)^k}{k!} = e^{2s-2}$

$\rightarrow \mathbb{E}[e^{tT} | N = n] = \mathbb{E}\left[e^{t\sum_{i=1}^{n} T_i} \,\middle|\, N = n\right] \overset{\text{Independence}}{=} \prod_{i=1}^{n} \mathbb{E}[e^{tT_i}] = (e^{e^t - 1})^n$

So, $\mathbb{E}[e^{tT} | N] = (e^{e^t - 1})^N$.

$\rightarrow \mathbb{E}[e^{tT}] = \mathbb{E}\left[\mathbb{E}[e^{tT} | N]\right] = \mathbb{E}\left[(e^{e^t - 1})^N\right] = e^{2(e^{e^t - 1}) - 2}$

Thus, $M_T(t) = e^{2(e^{e^t-1})-2}$

b) Remember that $M_X^{(n)}(0) = \mathbb{E}[X^n]$

$\rightarrow M_T'(t) = e^{2(e^{e^t-1})-2} \cdot 2e^{e^t-1} e^t$

$= 2e^{2e^{e^t-1}+e^t+t-3}$

$\rightarrow \mathbb{E}[T] = M_T'(0) = 2$

$\rightarrow M_T''(t) = 2e^{2e^{e^t-1}+e^t+t-3}\left(2e^{e^t-1}e^t + e^t + 1\right)$

$\rightarrow \mathbb{E}[T^2] = M_T''(0) = 2e^{2+1-3}\left(2+1+1\right) = 8$

$\rightarrow Var(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = 8 - 4 = 4$

<u>OR:</u> Since T_i's are iid Poisson(1), we see that $T \sim$ Poisson(N). So,

$\rightarrow \mathbb{E}[T] = \mathbb{E}\left[\mathbb{E}[T \mid N]\right] = \mathbb{E}[N] = 2$

$\rightarrow \mathbb{E}[T^2] = \mathbb{E}\left[\mathbb{E}[T^2 \mid N]\right] = \mathbb{E}[N + N^2] = 2 + 2 + 4 = 8$

$\rightarrow Var(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = 8 - 4 = 4$

MATH 505a QUALIFYING EXAM

September 20, 2004

You should try at least three problems; you may try all four.

1. There are three coins that show heads with probability 2/3, and tails otherwise. The first coin counts 8 points for a head and 3 for a tail, the second counts 5 points for both head and tail, and the third counts 4 points for a head and 10 for a tail. You and your opponent choose a coin, and you cannot choose the same coin. Each of you tosses your coin and the person with the larger score wins. Would you prefer to be the first to pick a coin, or the second?

2. A coin with probability p for heads is tossed n times. Find the expected number of "runs" and the variance of the number of "runs", where "run" is a sequence of identical outcomes. For example, the sequence $TTTHTHH$ has four runs, TTT, H, T and HH.

3. Let X_n be a sequence of independent identical distributed random variables with distribution function F such that $F(x) < 1$ for all x. Let $Y(y) = \min\{k : X_k > y\}$. Find the probability

$$P(Y(y) \leq E[Y(y)]).$$

Can you also show that the limit of this probability when $y \to \infty$ is $1 - 1/e$?

4. Let $X_1, X_2, \ldots, X_n$ be independent identically distributed normal random variables with mean μ and variance σ^2. Consider

$$S_n = X_1 + \cdots + X_n.$$

a) Find the moment generating function of S_n.

b) Find $E(e^{S_n})$.

c) What is the moment generating function of $(S_n - n\mu)/\sqrt{n}\sigma$?

Fall 2004, Applied Probability

1. Let B_i be the event that first person chooses i^{th} coin and A_i be the event that second person wins the game by choosing i^{th} coin.

→ $P(A_2 | B_1) = 3/9$, $P(A_3 | B_1) = 1/3 + 2/3 * 1/3 = 5/9$

So, if first person chooses first coin, second one has greater probability of winning by choosing third coin.

→ $P(A_1 | B_2) = 6/9$, $P(A_3 | B_2) = 3/9$

If first person chooses second coin, second person has greater probability of winning by choosing second coin.

→ $P(A_1 | B_3) = 2/3 * 2/3 = 4/9$, $P(A_2 | B_3) = 6/9$

If first person chooses third coin, second one has greater probability of winning by choosing second coin.

So, with the given probabilities and coins, there is no superior coin that makes the person who choses it the winner all the time. In this case, being second person to choose a coin is better, since the second person can make his/her choice to increase the probability of winning.

2. same as #1 of Spring 2006.

3. Firstly, we can see that $Y(y)$ is a nonnegative integer valued r.v. Then, consider that for $n \in \mathbb{N}$,

$$P(Y(y) = n) = P(X_1 \le y, X_2 \le y, \dots, X_{n-1} \le y, X_n > y)$$
$$= [F(y)]^{n-1}(1-F(y))$$) by independence.

We see that $Y(y) \sim \text{Geometric}(1-F(y))$. So, $E[Y(y)] = \frac{1}{1-F(y)}$. Then,

$$P(Y(y) \le E[Y(y)]) = P\left(Y(y) \le \frac{1}{1-F(y)}\right)$$
$$= \sum_{k=1}^{\lfloor \frac{1}{1-F(y)} \rfloor} [F(y)]^{k-1}(1-F(y))$$

$$= (1-F(y)) \frac{1-[F(y)]^{\lfloor \frac{1}{1-F(y)} \rfloor}}{1-F(y)}$$

$$= 1-[F(y)]^{\lfloor \frac{1}{1-F(y)} \rfloor}$$ is the desired probability.

Now consider

$$\lim_{y\to\infty} P(Y(y) \leq \mathbb{E}[Y(y)]) = 1 - \lim_{y\to\infty} F(y)^{\lfloor \frac{1}{1-F(y)} \rfloor}$$

$$= 1-\exp\left\{\lim_{y\to\infty}\left(\left\lfloor \frac{1}{1-F(y)} \right\rfloor \ln F(y)\right)\right\}$$

$$= 1-\exp\left\{\lim_{y\to\infty}\left(\frac{1}{1-F(y)} \ln F(y)\right)\right\}$$

L'Hôspitol's Rule

$$= 1-\exp\left\{\lim_{y\to\infty} \frac{1/F(y) \times f(y)}{-f(y)}\right\}$$

$$= 1-\exp\left\{\lim_{y\to\infty} -\frac{1}{F(y)}\right\}$$

$$= 1-e^{-1}$$

4 a) Let X be a $N(0,1)$ r.v. Consider

$$M_Z(t) = \mathbb{E}[e^{tZ}] = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = e^{t^2/2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz}_{=1\text{, since the integrand is pdf of } N(t,1) \text{ r.v.}} = e^{t^2/2}$$

Then, since $X_i = \mu + \sigma Z$, we have

$$M_{X_i}(t) = \mathbb{E}[e^{t\mu+\sigma t Z}] = e^{t\mu}\,\mathbb{E}[e^{(\sigma t)Z}] = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

Then,

$$M_{S_n}(t) = \mathbb{E}[e^{tS_n}] = \mathbb{E}[e^{X_1 t + \cdots X_n t}] \overset{\text{by independence}}{=} \prod_{i=1}^{n} \mathbb{E}[e^{X_i t}] = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$$= e^{n\mu t + \frac{n\sigma^2 t^2}{2}}$$

b) $\mathbb{E}[e^{S_n}] = M_{S_n}(1) = e^{n\mu + \frac{n\sigma^2}{2}}$.

c) Let us denote $\frac{S_n - n\mu}{\sqrt{n}\sigma}$ by T_n. Consider

$$M_{T_n}(t) = \mathbb{E}\left[e^{tT_n}\right]$$

$$= \mathbb{E}\left[e^{\frac{tS_n - tn\mu}{\sqrt{n}\sigma}}\right]$$

$$= e^{\frac{-tn\mu}{\sqrt{n}\sigma}} \cdot \mathbb{E}\left[e^{(t/\sqrt{n}\sigma)S_n}\right]$$

$$= e^{(-\sqrt{n}\mu/\sigma)t} \cdot e^{(\sqrt{n}\mu/\sigma)t} \, e^{t^2/2}$$

$$= e^{t^2/2}$$