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

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

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

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

Thus,

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

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

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

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

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

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