The solutions which we reproduce below are far more elaborate than what was expected on the final. After all, not many students have laser printers and Mathematica running on a laptop during the exam! Our purpose is to give as full an explanation as possible, so *you* understand the solution.

Problem 1. (15 points) A certain investor is interested in a group of 30 stocks. Among these stocks are 17 stocks traded at the New York Stock Exchange (NYSE) and 13 stocks traded at the American Stock Exchange (Amex): Among the NYSE stocks, the prices of 12 stocks are up since the previous trading day and the prices of 5 stocks are down. Among the Amex stocks, the prices of 7 stocks are up since the previous trading day and the prices of 6 stocks are down. When appropriate, you may leave your answer in terms of factorials.

(a) Suppose one stock is selected at random. Let A represent the event that a stock traded on the NYSE is chosen, and let B represent the event that a stock whose price has risen is chosen. Find $P(A)$ and $P(A|B)$.
(b) Are A and B independent? Show your work.
(c) The investor needs to read a report on each of the Amex stocks whose price is up since the previous trading day, and must choose the order in which to read them. How many ways are there for the investor to do this? Show your work.
(d) How many ways can the investor choose a group of 5 different stocks to buy from the group of 13 Amex stocks?
(e) Suppose that the investor wants to invest in some of the 13 Amex stocks. How many ways can the investor choose a group of 3 Amex stocks that are up and a group of 2 Amex stocks that are down?

Solution.

(a) There are a total of 30 stocks; the event A represents choosing one of the 17 NYSE stocks, so $P(A) = 17/30 = 0.5667$, the number of successes over the number of possibilities.

By definition, $P(A|B) = P(A \text{ and } B)/P(B)$, where B is the event "the price has risen". There are 19 stocks which are up, of which 12 are NYSE stocks; thus $P(A \text{ and } B) = 12/30$, and $P(B) = 19/30$. Therefore $P(A|B) = 12/19 = 0.6316$. Of course, this has also the common-sensical solution: probabilities *given* B are over a new universe of 19 simple events (the stocks that went up), and 12 of them were from the NYSE, thus the probability of choosing an NYSE stock from the 19 which went up is $12/19$.

(b) For A and B to be independent, we must have

$$P(A \text{ and } B) = P(A)P(B)$$

(because that's the *definition* of independent events). So we are asking whether

$$\frac{12}{19} = \frac{17}{30} \times \frac{19}{30},$$

and this is certainly false. (For one thing, the numerator on the right side isn't divisible by 3, and the numerator on the left is. Or pull out your calculator and check.)

(c) 7 of the Amex stocks are up, and you're asked how many ways the investor can order these 7 stocks. The answer is

$$7! = 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 = 5040$$

Perhaps he should quit daydreaming and read them alphabetically `;-)`

(d) This is $\binom{13}{5}$ (which is even *pronounced* "13 choose 5"), or

$$\binom{13}{5} = \frac{13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 1287$$

ways. (You can also use

$$\binom{13}{5} = \frac{13!}{5! \cdot 8!} = \frac{6227020800}{120 \cdot 40320}$$

if you have a lot of time to waste.)

(e) There are 7 Amex stocks that are up, and the investor wants to choose 3 of them; there are 6 Amex stocks which went down and he wants to choose 2 of these.

Now there are $\binom{7}{3} = 35$ ways of choosing the stocks which went up, and $\binom{6}{2} = 15$ ways of choosing the stocks which went down. By the multiplication principle, there are $35 \times 15 = 525$ ways of doing *both* these things.

□

Problem 2. (15 points) USC students are either in the college of Letters, Arts and Sciences (LAS), the graduate school (GS) or the professional schools (PS). The percentage of students in each group that support Al Gore for president in the 2000 election are as follows: LAS, 30%; GS, 40%; PS, 20%. There are currently 12000 students enrolled in LAS, 6000 students enrolled in the graduate school, and 9000 students enrolled in the professional schools.

(a) What is the probability that a student chosen at random from the USC student body is enrolled in one of the professional schools?
(b) What is the probability that a student chosen at random from the USC student body will support Al Gore in the 2000 presidential election?
(c) If you pass a student on campus and he or she is wearing a *Gore in 2000* button, what is the probability that the student is enrolled in LAS?

Solution. □

(a) There are a total of $12000 + 6000 + 9000 = 27000$ students enrolled at USC, of whom 9000 are enrolled in PS; thus the probability that a randomly-chosen student at USC is in PS is $9000/27000 = 1/3$.

(b) We compute the number of students in each school who support Gore:

School	Enrolled	Percentage	For Gore
LAS	12000	30%	3600
GS	6000	40%	2400
PS	9000	20%	1800
Total	27000		7800

Thus 7800 of the enrolled 27000 students are in favor of Gore, for a probability of $7800/27000 = 0.2889$.

(c) Assuming that a student wearing a *Gore in 2000* button is a Gore supporter, the correct answer is $3600/7800 = 0.4615$: the population is restricted to Gore supporters (the last column of the preceding table) and 3600 of the 7800 students who are Gore supporters are in LAS.

□

Problem 3. (15 points) The manager of a certain small restaurant has observed that the most popular entree is the fish entree. 30% of the diners select the fish entree.

(a) Find the expected number of fish entrees ordered by a party of 4 diners.
(b) Find the standard deviation in the number of fish entrees ordered by a party of 4 diners.
(c) Find the probability that in a party of 4 diners, exactly 2 of them order the fish entree.

Solution. This can be treated as a binomial probability problem: an individual diner will select the fish entree with probability $p = 0.30$, or *not* select the fish entree, with probability $1 - p = 0.7$.

(a) Assuming the four diners make their choices independently, the answer is $np = 4 \times 0.3 = 1.2$.
(b) Since this is a binomial probability, the standard deviation is $\sigma\sqrt{n}$, where $n = 4$ and $\sigma = \sqrt{p(1-p)} = \sqrt{0.3 \times 0.7} = 0.4583$ is the standard deviation of a single Bernoulli trial. The answer is therefore $2 \times 0.4583 = 0.9165$.
(c) The exact binomial probability is

$$\binom{4}{2}(0.3)^2(0.7)^2 = 6 \times 0.09 \times 0.49 = 0.2646$$

Problem 4. (20 points) An office furniture store is having a sale on a certain type of office chair. The chairs are priced at \$80 per chair. The probability distribution for the number of chairs sold to an individual customer is

Number of chairs	1	2	3	4
Probability	0.15	0.40	0.10	0.35

(a) Find the probability that the number of chairs sold to an individual customer is odd.
(b) Find the expected number of chairs sold to an individual customer.
(c) Find the standard deviation in the number of chairs sold to an individual customer.
(d) The delivery charge for an order of chairs is \$50. Find the expected amount a customer pays for the chairs he or she buys, *including* the delivery charge.
(e) Find the standard deviation in the amount a customer pays for the chairs he or she buys, including the delivery charge.

Solution. Let X denote the number of chairs sold to a customer.

(a) The event "X is odd" is the same as "$X = 1$ or $X = 3$", and these are mutually exclusive; thus

$$P(X \text{ is odd}) = P(X = 1) + P(X = 3) = 0.15 + 0.10 = 0.25$$

(b) The expected value is obtained by multiplying the number of chairs sold, by the probability of selling that many chairs; or

$$E(X) = 1 \times 0.15 + 2 \times 0.40 + 3 \times 0.10 + 4 \times 0.35 = 2.65$$

(c) The standard deviation is defined to be the square root of the variance, which is

$$\mathrm{Var}(X) = \sum_{i=1}^{4} (x_i - \mu)^2 P(X = x_i), \qquad \mu = E(X).$$

This is most easily done by hand by extending the table by three rows, then adding the entries in the last row.

x_i	1	2	3	4
$P(X = x_i)$	0.15	0.40	0.10	0.35
$x_i - \mu$	−1.65	−0.65	0.35	1.35
$(x_i - \mu)^2$	2.7225	0.4225	0.1225	1.8225
$(x_i - \mu)^2 P(X = x_i)$	0.408375	0.169	0.01225	0.637875

The result is 1.2275, which is the *variance*, remember; the standard deviation is the square root of this, 1.10793. *There is no question of dividing by $n-1$ instead of n in this problem;* this is supposed to be the exact probability distribution; *nothing* is being deduced from *data*.

(d) The customer cost is $C = 80X + 50$, because each chair costs \$80 and there's a \$50 delivery charge which doesn't depend on the number of chairs delivered. Thus

$$E(C) = 80E(X) + 50 = 80 \times 2.65 + 50 = \$262.$$

(Well, duh! The customer buys 2.65 chairs on average, at \$80 each, and pays \$50 extra.)

(e) This is less of a "Duh!" question. The formula for the standard deviation of $C = 80X + 50$ is $\sigma_C = 80\sigma_X$, which for this problem works out to be $80 \times 1.10793 = \$88.63$. In general, when finding the standard deviation of $aX+b$, remember that a *shift* (constant added) doesn't change the standard deviation (which measures the *spread* of data about the mean); shifting simply moves everything over, it doesn't change the spreading tendencies. Also remember that *scaling* a variable multiplies the standard deviation by the absolute value of the scale. If we measure something in feet, and obtain a standard deviation of 2 feet, then when we convert to inches we will get a standard deviation of 24 inches.

□

Problem 5. (15 points) Let X be a continuous random variable denoting the time required for an assembly operation. It is known that X has a uniform distribution between 30 and 40 seconds.

(a) Determine the proportion of assemblies that require more than 37 seconds.
(b) What assembly time is exceeded by 90% of the assemblies?
(c) Determine the mean and standard deviation of the assembly time.

Solution. We show the graph of the uniform distribution in Figure 1. Since the area has to be 1, and the base of the rectangle is 10, it follows that the height of the shaded rectangle is 0.1.

(a) Refer to Figure 2. The area of the darker rectangle is the proportion of assemblies that require more than 37 seconds; it's obviously 0.3, because its base is 3 (from 37 to 40) and its height is 0.1.
(b) Refer to Figure 3. The darker area represents the 90% of the assemblies which take the most time; since the height is 0.1, therefore the base of this rectangle must be 9. This means that such assemblies take ≥ 31 seconds to complete.

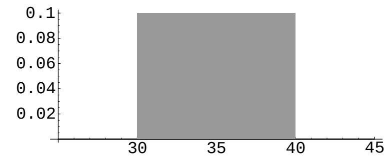


FIGURE 1. Uniform density distribution on interval [30, 40]

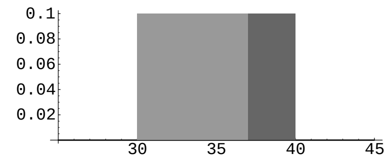


FIGURE 2. Uniform density with $x \geq 37$ darkened

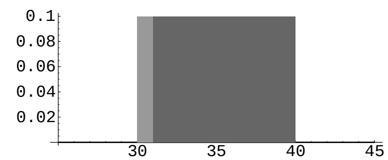


FIGURE 3. Uniform density with $x \geq 31$ (top 90%) darkened

(c) Finding the expected value can be done without any calculus at all; it represents the "balancing line" for the graph in Figure 1, the vertical line along which the graph would balance if the line were a knife edge. This obviously happens at $\mu = 35$ seconds, midway between 30 and 40 seconds.

The standard deviation must be calculated (or you must use a formula for the uniform distribution). The formula is: for a random variable uniformly distributed on an interval $[a, b]$, the variance is $(b-a)^2/12$ and the standard deviation is therefore

$$\frac{b-a}{2\sqrt{3}}.$$

So in this case the standard deviation is $10/(2\sqrt{3}) = 2.887$ seconds. You can also compute it as

$$\mathrm{Var}(E) = \int_{30}^{40} 0.1(x-35)^2\,dx = 8.3333,$$

from which it follows that $\sigma = \sqrt{8.3333} = 2.887$.

$\square$

Problem 6. (20 points) The speeds (in mph) of cars on Main Street are observed to be normally distributed with mean 37 and standard deviation 4.5. The speed limit is 35 mph.

(a) What fraction of drivers are exceeding the speed limit?
(b) Officer Krupke decides to give tickets only to the fastest 10% of cars on Main Street. How fast must a driver go to get a ticket from Officer Krupke?
(c) Under the assumptions in (b), find the approximate probability that out of the next 150 cars on Main Street, at most 10 receive tickets from Officer Krupke.

Solution. Let X denote the speed of a car. Then

$$Z = \frac{X - 37}{4.5}$$

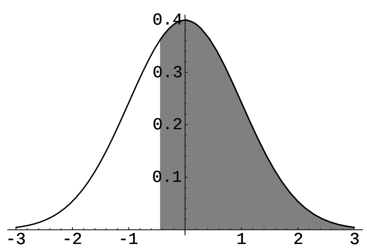
is normally distributed with mean 0 and standard deviation 1.

(a) We are asked to compute $P(X > 35)$, and we do so by parallelling the construction of Z from X:

$$\begin{aligned} P(X > 35) &= P(X - 37 > -2) \\ &= P\left(\frac{X-37}{4.5} > \frac{-2}{4.5}\right) \\ &= P(Z > -0.444) \\ &= 0.5 + P(0 \le Z \le 0.444) \\ &= 0.6715. \end{aligned}$$

Thus about 67% of the drivers are speeding.

The situation is summarized in the following graph:



(b) The top 10% of speeders are obtained by computing the cutoff $Z_{.1} = 1.282$; that is, $P(Z \ge 1.282) = 0.1$. Thus

$$\begin{aligned} 0.1 &= P\left(\frac{X-37}{4.5} \ge 1.282\right) \\ &= P(X \ge 37 + 4.5 \cdot 1.282) \\ &= P(X \ge 42.769) \end{aligned}$$

Thus the 10% fastest speeders are those travelling about 42.8 mph or faster.

(c) This is a binomial probability problem, with $n = 150$ and $p = 0.1$. If X is the total number of the 150 cars which will be stopped by Officer Krupke, then we are asking for $P(X \le 10)$. The exact answer is

$$\sum_{x=0}^{10} \binom{150}{x} 0.1^x 0.9^{150-x},$$

but this is difficult to compute without a computer (or at least a good calculator). If you have one, the exact answer turns out to be 0.105963. Hold this in mind for a little while.

Without a computer, we can do this by approximating by a normal distribution with mean $\mu = np = 150 \times 0.1 = 15$ and standard deviation

$$\sigma = \sqrt{np(1-p)} = \sqrt{150 \times 0.1 \times 0.9} = 3.67423.$$

We take $Z = (X - 15)/3.67423$. But to find $P(X \le 10)$ we use the continuity correction, and instead compute $P(X \le 10.5)$:

$$\begin{aligned} P(X \le 10.5) &= P(X - 15 \le -4.5) \\ &= P\left(\frac{X-15}{3.67423}\right) \le \frac{-4.5}{3.67423}) \\ &= P(Z \le -1.22) \\ &= 0.1112. \end{aligned}$$

This isn't very far off from the exact answer 0.105963.

$\square$

Problem 7. (15 points) A doctor working in a clinic finds that the amount of time he spends with a patient has a distribution with mean 5 minutes and standard deviation 2 minutes. One day he finds that there are 30 patients in the waiting room. Assuming that he sees the patients consecutively with no gaps between them starting at 10 a.m., what is the probability that the doctor will be free for lunch at 12:15 p.m.?

Solution. Let X_1 be the time the doctor spends with the first patient, X_2 the time he spends with the second, etc. Then the total time he spends with the patients is

$$X = X_1 + X_2 + \cdots + X_{30}.$$

Since there are 135 minutes between 10:00 AM and 12:15 PM, we are being asked for $P(X \leq 135)$.

Assuming the times he spends with patients are independent, we have

$$E(X) = nE(X_i) = 30 \cdot 5 = 150 \text{ minutes},$$

and

$$\sigma_X = \sigma_{X_i}\sqrt{n} = 2\sqrt{30} = 10.9545 \text{ minutes}.$$

Since the X_i are normally distributed, so is X, and therefore

$$Z = \frac{X - 150}{10.9545}$$

has the standard normal distribution. We compute the desired $P(X \leq 135)$ by parallelling the construction of Z from X:

$$\begin{aligned} P(X \leq 135) &= P(X - 150 \leq -15) \\ &= P\left(\frac{X-150}{10.9545} \leq \frac{-15}{10.9545}\right) \\ &= P(Z \leq -1.3693) \end{aligned}$$

From the Z-table we have $P(0 \leq Z \leq 1.37) = 0.4131$; we must subtract this from 0.5, for a total of 0.0869. Not much chance the doc will get his lunch on time. □

Problem 8. (25 points) The Santa Berenica city council contracted a statistical consultant to prepare a study about the impact of abolishing rent control regulations. As a preliminary step, the researcher decided to estimate the mean rent increase for a one-bedroom apartment in Santa Berenica. From a random sample of 5 one-bedroom apartments in the area he obtained the following monthly rent increases (in dollars): 130, 220, 175, 140, 100.

(a) Find point estimates for the mean and standard deviation of the rent increase for a one-bedroom apartment in Santa Berenica.
(b) Determine a 95% confidence interval for the mean rent increase for one-bedroom apartments in Santa Berenica. (If you need additional assumptions about the rent increases, use them and list them in the next question.)
(c) What assumptions, if any, about the apartment rent increases did you use to answer the previous question?
(d) After the council meeting the consultant was asked to prepare a more detailed study estimating the mean apartment rent increase within ±\$10 with confidence of 99%. How many apartments does he need to select in total (including the original 5) to obtain a confidence interval with the required width?

Solution. Since this is a sampling problem, we must compute the mean and *sample* standard deviation (you remember, dividing by $n-1$ instead of by n). Also, since the sample size is small we'll have to use the Student t-distribution.

(a) We compute

$$\overline{X} = \frac{130 + 220 + 175 + 140 + 100}{5} = \$153$$

and

$$\begin{aligned} s^2 = \frac{1}{4}\big((130-153)^2 &+ (220-153)^2 + (175-153)^2 \\ &+ (140-153)^2 + (100-153)^2\big) = 2120, \end{aligned}$$

from which $s = 46.0435$.

(b) For $\alpha = 0.05$ we compute $t_{4,\alpha/2} = 2.776$, so that the 95% confidence interval has endpoints

$$\overline{X} \pm t_{4,\alpha/2}\frac{s}{\sqrt{n}} = 153 \pm 2.776 \times \frac{46.0435}{\sqrt{5}} = 153 \pm 57.1614.$$

Thus the confidence interval is approximately [95.84, 210.16].

(c) We had to assume a normal distribution for rent increases, since $n = 5$ is too small for the Central Limit Theorem to have much effect. We also had to assume the sample was random.

(d) There is a formula for the value of n which you probably wrote down on the page you were allowed to bring into the exam,

$$n = \frac{4t^2_{\alpha/2,n-1}s^2}{w^2},$$

but I'm not much of a fan of memorizing formulas. Instead, I favor remembering the confidence interval,

$$\overline{X} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}.$$

Since we're required to find the mean apartment rent increase within ±\$10, this means we should set

$$t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} = 10.$$

But we don't know n, so how are we supposed to know $t_{\alpha/2,n-1}$?! The answer is, for $t_{\alpha/2,n-1}$ we use the *old* value of n: $t_{0.005,4} = 4.604$. Thus the equation becomes

$$4.604\frac{46.0435}{\sqrt{n}} = 10,$$

which means we should have $\sqrt{n} = 21.1984$, so that $n = 449.373$. We have to take $n = 450$ to accomplish this. (This is a very conservative number.)

□

Problem 9. (20 points) Susie, an 8-year-old budding statistician, observes that whenever her buttered toast falls on the floor, it usually seems to land butter-side-down. (It is buttered on one side.) One day, while her Mom is out, Susie drops her buttered toast on the floor 40 times, and she observes that it lands butter-side-down 27 times.

(a) Find a point estimate for the proportion of time that the toast lands butter-side-down. State whether the estimator you are using is biased or unbiased.
(b) Find Susie's 90% confidence interval for the proportion of time that the toast lands butter-side-down.
(c) Susie's 10-year-old brother Mike is unimpressed—he tells Susie he will determine the probability of butter-side-down to within ±0.02, with 90% confidence, and will use Susie's data as a mere preliminary sample. How many more times (in addition to Susie's 40) must Mike drop the toast, to achieve his desired accuracy?

Solution. This is a binomial probability problem, with $n = 40$ and the observed $\hat{p} = 27/40 = 0.675$.

(a) The point estimate for the proportion is $\hat{p} = 0.675$, and it is unbiased (which means that its expectation is the true proportion).
(b) We'll approximate the binomial probability as a normal probability—since $n\hat{p}(1-\hat{p}) = 8.775 > 5$, this satisfies the rule of thumb as to when we can do this.

The 90% confidence interval is therefore given by

$$\hat{p} \pm Z_{.05}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.675 \pm 1.645\sqrt{\frac{0.219375}{40}} \qquad = 0.675 \pm 0.1218$$

i.e. $[0.5532, 0.7968]$.

(c) The confidence interval is $\hat{p} \pm Z_{.05}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$, and thus we want

$$1.645\frac{0.468375}{\sqrt{n}} = 0.02.$$

The solution of this is

$$n = \left(\frac{1.645 \times 0.468375}{0.02}\right)^2,$$

i.e. $n = 1484.09$. Thus Mike will need a total of 1485 measurements; since Susie has already made 40, that means he will need to make 1445 more measurements. A rather dirty floor, when Mom comes home . . .

□

Problem 10. (20 points) A researcher for Red Cars of America wants to use a random sample of 426 tickets to see if there is strong evidence to support the idea that an unusually high percentage of tickets go to drivers of red cars. Approximately 11% of the cars on California's roads are red in color.

(a) Formulate appropriate null and alternative hypotheses for the researcher to use.
(b) Choose an appropriate test statistic and rejection rule to test your null hypothesis against your alternative hypothesis at the 5% level of significance.
(c) During a recent 24 hour period, on a particular stretch of freeway, the Highway Patrol issued 426 tickets, of which 63 went to drivers of red cars. Determine whether or not the null hypothesis should be rejected at the 5% level of significance. Explain your answer.
(d) Find the P-value for the data in this test. Your answer may either be an exact value, or else the smallest interval which can be determined from the table provided.
(e) Circle **T**rue or **F**alse: as a result of this hypothesis test, the researcher can conclude:

There is not strong evidence that an unusually high percentage of tickets go to red cars.	T F
Red cars apparently receive an unusually high percentage of tickets.	T F
Drivers of red cars are treated unfairly by the police.	T F

Solution. This is a binomial probability problem, but the value of n is so large that we can approximate by the standard normal distribution Z.

(a) The null hypothesis is: $H_0 : p = p_0$ (where p is the proportion of the tickets which were given to red cars, and $p_0 = 0.11$); the alternative hypothesis is $p > p_0$ (that red cars get *more* tickets than expected).

(b) We form the test statistic

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{\hat{p} - 0.11}{0.31289/\sqrt{n}}$$

The 5% significance level corresponds to Z-cutoff of $Z_{0.05} = 1.645$, thus we reject H_0 if $Z > 1.645$.

(c) We compute $\hat{p} = 63/426 = 0.147887$ and $0.31289/\sqrt{426} = 0.0151596$, and thus

$$Z = \frac{0.147887 - 0.11}{0.0151596} = 2.499.$$

Since this value of Z is indeed > 1.645, we *reject* the null hypothesis in favor of the alternative hypothesis.

(d) The P-value is $P(Z > 2.499) = 0.0062$ (which you calculate by subtracting the value in the Z-table corresponding to $Z = 2.5$ from 0.5).

(e) There certainly *is* strong evidence that an unusually high percentage of tickets goes to red cars, so you should have circled "F". It is also true that red cars apparently receive an unusually high percentage of tickets, so you should have circled "T" for this statement. But you *cannot* conclude that "Drivers of red cars are treated unfairly by police"; nothing in the experiment measured fairness (which would be difficult to quantify in any case); perhaps drivers of red cars drive faster!

□

Problem 11. (20 points) FAA regulations state that the cables manufactured for use in aircraft must have a mean length of exactly 40 inches. The FAA investigates a random sample of 15 aircraft cables made by AirTrain, Inc. to test whether the cables meet regulations. If they are either too long or too short, on average, then AirTrain may be fined. Assume that the cable lengths are normally distributed.

(a) Formulate the appropriate null and alternative hypothesis that the FAA investigators should use.

(b) Choose a suitable test statistic and rejection rule to test your null hypothesis against your alternative hypothesis at the 1% level of significance.

(c) In the sample of 15 cables, the average cable length was found to be 39.7 inches, with a sample standard deviation of 0.4 inches. Determine whether the null hypothesis should be rejected at the 1% level of significance. Explain your reasoning.

(d) Find the P-value for the data in this test. Your answer may either be an exact value, or else the smallest interval which can be determined from the table provided.

(e) A second investigation of cables made by another company resulted in a P-value of 0.028. At which of the following significance levels should the null hypothesis be rejected for this second investigation? Circle all that apply.

0.5% 1% 2.5% 5% 10%

Solution.

(a) Null hypothesis H_0: $\mu = 40$, where μ is the length. Alternative hypothesis: $H_a : \mu \neq 40$. This is a two-sided hypothesis, since cables being too short *or* too long will be out-of-spec.

(b) Since we have 15 data measurements, we use the student t-test with $\nu = 14$ degrees of freedom:

$$t_{14} = \frac{\overline{X} - 40}{s/\sqrt{15}},$$

where s is the sample standard deviation. For $\alpha = 0.01$, we set up as the rejection rule

$$|t| > t_{\alpha/2,14} = 2.977.$$

That is, we *reject* H_0 if $t < -2.977$ or $t > 2.977$.

(c) So we compute

$$t = \frac{\overline{X} - 40}{s/\sqrt{15}} = \frac{39.7 - 40}{0.4/\sqrt{15}} = -2.90474.$$

Since $|t|$ is *not* greater than 2.977, we do *not* reject H_0 at the 1% level of significance.

(d) We need to compute $P(|t| > 2.90474)$. This is hard to do without a smart calculator or a computer; if we *do* have one available, we find

$$P(t < -2.90474) = P(t > 2.90474) = 0.00576675,$$

and therefore $P(|t| > 2.90474) = 2 \times 0.00576675 = 0.0115335$.

Otherwise, we must look at the tables of the t cutoffs for 14 degrees of freedom. These are

ν	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$
14	1.345	1.761	2.145	2.624	2.977

The computed value 2.90474 lies between the last two entries of this table, corresponding to a value between $t_{.01}$ and $t_{.005}$. Be careful! If you answered "Between $\alpha = 0.01$ and $\alpha = 0.005$", you forgot that this is a *two-sided* test; you should have said, "Between $\alpha/2 = 0.01$ and $\alpha/2 = 0.005$", which corresponds to between $\alpha = 0.02$ and $\alpha = 0.01$. You can also try linear interpolation; this consists of visualizing the t cutoffs as lying on a straight line, with the two known points $(0.01, 2.624)$ and $(0.005, 2.977)$ lying on the line; we then want to know the x coordinate such that $(x, 2.90474)$ lies on the line. Since 2.90474 is 79.5% of the way from 2.624 to 2.977, the correct value of x is 79.5% of the way from 0.01 to 0.005 (this is *decreasing*), or 0.00602351. Remembering to double it (α instead of $\alpha/2$, remember?), we get a P-value of 0.012047, which doesn't compare badly to the actual value of 0.0115335. Linear interpolation is crude, but it's a lot better than saying "Somewhere between 1% and 2%".

The strategy of taking an exam virtually requires you to say "Somewhere between 1% and 2%", unless you finish the exam early and want to try to impress the professors.

(e) Remember that the P-value represents the probability of a result *as extreme as, or more extreme than,* the measured value. So the other company's P-value of 2.8% means that we should reject the null hypothesis for any level of significance *greater* than 2.8%; i.e. at the 5% and 10% levels. We cannot reject it at the other levels, because we do not satisfy those significance levels.

□