Solutions to Sample Problems in Probability and Statistics

1) a) $26 * 26 * 26 * 26 * 26 * 10 = 118,813,760$

   b) $26 * 26 * 26 * 26 * 26 * 5 = 59,406,880$

   c) $25 * 25 * 25 * 25 * 25 * 10 = 97,656,250$

2) Let A be the event that the sample contains no more than 2 defective units. Then,

$$\begin{aligned}\mathrm{P}[A] &= \frac{\binom{80}{4}\binom{20}{0} + \binom{80}{3}\binom{20}{1} + \binom{80}{2}\binom{20}{2}}{\binom{100}{4}} \\ &= \frac{3,825,180}{3,921,225} = \frac{255,012}{261,415} \\ &= 0.976\end{aligned}$$

3) Let A be the event that the complaint was product appearance.
   Let D be the event that the complaint originated during the guarantee period. Then,

$$\mathrm{P}\left[A \mid D\right] = \frac{\mathrm{P}[A \cap D]}{\mathrm{P}[D]} = \frac{0.32}{0.63} = \frac{32}{63} = 0.51$$

4) Let M be the event that the main engine is operable.
   Let B be the event that the backup engine is operable.
   Given $\mathrm{P}[M] = 0.95$, $\mathrm{P}[B] = 0.80$, and $\mathrm{P}[M \cup B] = 0.99$, then

   a) $\mathrm{P}[M \cap B] = \mathrm{P}[M] + \mathrm{P}[B] - \mathrm{P}[M \cup B] = 0.95 + 0.80 - 0.99 = 0.76$

   b) $\mathrm{P}[M^c \cap B] = 0.80 - 0.76 = 0.04$

   c) $\mathrm{P}[B^c \cap M] = 0.95 - 0.76 = 0.19$

   d) $\mathrm{P}[M^c \cap B^c] = \mathrm{P}[(M \cup B)^c] = 1 - \mathrm{P}[M \cup B] = 1 - 0.99 = 0.01$

   e) $\mathrm{P}\left[B \mid M^c\right] = \frac{\mathrm{P}\left[B \cap M^c\right]}{\mathrm{P}[M^c]} = \frac{0.04}{0.05} = 0.80$

5) Let E be the event that the shutdown is due to equipment failure.
   Let O be the event that the shutdown is due to operator error.
   Given: $\mathrm{P}[E \cap O^c] = 0.10$, $\mathrm{P}[E \cap O] = 0.05$, $\mathrm{P}[O] = 0.40$.

   a) $\mathrm{P}[E \cup O] = 0.10 + 0.05 + 0.35 = 0.50$ from Venn Diagram

   b) $\mathrm{P}[O \cap E^c] = 0.35$

   c) $\mathrm{P}[O^c \cap E^c] = \mathrm{P}[(O \cup E)^c] = 1 - \mathrm{P}[O \cup E] = 1 - 0.5 = 0.5$

   d) $\mathrm{P}\left[0 \mid E\right] = \frac{\mathrm{P}[O \cap E]}{\mathrm{P}[E]} = \frac{0.05}{0.15} = 0.33\dot{3}$

   e) $\mathrm{P}\left[0 \mid E^c\right] = \frac{\mathrm{P}\left[O \cap E^c\right]}{\mathrm{P}[E^c]} = \frac{0.35}{0.85} = 0.4118$

6) Let C be the event that the carpenters go on strike.
   Let P be the event that the plumbers go on strike.
   Let D be the event that there is a delay in project completion.
   Given: $\mathrm{P}\left[D \mid C \cap P\right] = 1$, $\mathrm{P}\left[D \mid C \cap P^c\right] = 0.8$, $\mathrm{P}\left[D \mid C^c \cap P\right] = 0.4$, $\mathrm{P}\left[D \mid C^c \cap P^c\right] = 0.05$, $\mathrm{P}\left[P \mid C\right] = 0.6$, $\mathrm{P}\left[C \mid P\right] = 0.30$, $\mathrm{P}[P] = 0.10$

   a) $\mathrm{P}[P \cap C] = \mathrm{P}[P]\,\mathrm{P}\left[C \mid P\right] = 0.1(0.3) = 0.03$

   b) $\mathrm{P}[C] = \frac{\mathrm{P}[C \cap P]}{\mathrm{P}\left[P \mid C\right]} = \frac{0.03}{0.6} = 0.05$

c) $\text{P}[P^c \cap C] = \text{P}[C]\,\text{P}\left[P^c \mid C\right] = \text{P}[C]\,(1 - \text{P}\left[P \mid C\right]) = 0.05(1-0.6) = 0.02$

d) $\text{P}[C^c \cap P] = \text{P}[P]\,\text{P}\left[C^c \mid P\right] = \text{P}[P]\,(1 - \text{P}\left[C \mid P\right]) = 0.1(1-0.3) = 0.07$

e) $\text{P}[C^c \cap P^c] = 1 - 0.03 - 0.02 - 0.07 = 0.88$

f) $\text{P}[D] = 0.03(1) + 0.02(0.8) + 0.07(0.4) + 0.88(0.05) = 0.118$

g)

i) $\text{P}\left[P \cap C \mid D\right] = \frac{0.03(1)}{0.118} = 0.254$

ii) $\text{P}\left[C \cap P^c \mid D\right] = \frac{0.02(0.8)}{0.118} = 0.136$

iii) $\text{P}\left[C \mid D\right] = \text{P}\left[(C \cap P) \cup (C \cap P^c) \mid D\right] = \frac{0.03(1)+0.02(0.8)}{0.118} = 0.390$

7) Given: $\text{E}[X] = 2$, $\text{E}\left[X^2\right] = 29$, $\text{E}[Y] = 4$, and $\text{E}\left[Y^2\right] = 52$.
X and Y are independent.
$W = X + Y$ and $Z = 2X$
Thus,

$$
\begin{aligned}
\text{Var}[X] &= \text{E}\left[X^2\right] - \text{E}^2[X] = 29 - 2^2 = 25 \\
\text{Var}[Y] &= \text{E}\left[Y^2\right] - \text{E}^2[Y] = 52 - 4^2 = 36 \\
\text{E}[W] &= \text{E}[X+Y] = 2 + 4 = 6 \\
\text{Var}[W] &= \text{Var}[X] + \text{Var}[Y] = 25 + 36 = 61 \\
\text{E}[Z] &= \text{E}[2X] = 2(2) = 4 \\
\text{Var}[Z] &= 4(25) = 100 \\
\text{Cov}[W,Z] &= \text{E}[WZ] - \text{E}[W]\,\text{E}[Z] \\
\text{E}[WZ] &= \text{E}[(X+Y)(2X)] = \text{E}\left[2X^2 + 2XY\right] = 2\text{E}\left[X^2\right] + 2\text{E}[X]\,\text{E}[Y] = 2(29) + 2(2)(4) = 74 \\
\text{Cov}[W,Z] &= 74 - 6(4) = 50 \\
\rho_{WZ} &= \frac{50}{\sqrt{61}\sqrt{100}} = 0.64
\end{aligned}
$$

8) a) Let N_{10} be the number of floods in a 10-year period.
N_{10} follows a Poisson distribution with $\lambda = 1/8$ floods per year. The probability of zero floods in 10 years is thus

$$\text{P}[N_{10} = 0] = \frac{(10/8)^0}{0!}e^{-10/8} = 0.2865$$

b) Let N_{15} be the number of floods in a 15-year period.
N_{15} follows a Poisson distribution with $\lambda = 1/8$ floods per year. The probability of no more than 3 floods in 15 years is thus,

$$
\begin{aligned}
\text{P}[N_{15} \leq 3] &= \text{P}[N_{15} = 0] + \text{P}[N_{15} = 1] + \text{P}[N_{15} = 2] + \text{P}[N_{15} = 3] \\
&= \frac{(15/8)^0}{0!}e^{-15/8} + \cdots + \frac{(15/8)^3}{3!}e^{-15/8} \\
&= 0.879
\end{aligned}
$$

Thus, $\text{P}[X > 3] = 1 - 0.879 = 0.121$

9) Let X be the daily water consumption of a city.
X is normally distributed with mean $\mu = 500,000$ and standard deviation $\sigma = 150,000$ gpd. Then

a) $$\begin{aligned} \text{P}[X > 750,000] &= \text{P}\left[\frac{X-\mu}{\sigma} > \frac{750,000 - 500,000}{150,000}\right] \\ &= \text{P}[Z > 1.67] \\ &= 0.0475 \end{aligned}$$

b) $$\begin{aligned} \text{P}[Z > 2.33] = .01 &\Rightarrow \text{P}[Z\sigma + \mu > 2.33(150,000) + 500,000] = .01 \\ &\Rightarrow \text{P}[X > 849,500] = .01 \end{aligned}$$

So the water supply must be at least 849,500 gallons in any given day.

10) Minitab was used to do some of the calculations and the plots. Other statistical software packages would do as well. These calculations might also be done using a calculator.

a) One must assume that the lengths of the spanner bushings are normally distributed. When one looks at a histogram of the sampled data (see below), this does not appear to be the case. Granted the sample size is small, but the validity of the assumption is suspect. Therefore, the results are to be viewed with more than the usual amount of skepticism.

```
NOTE Question #10 - Part I: Length Measurements
MTB > read c1
DATA> 1.1375
DATA> 1.1390
DATA> 1.1420
DATA> 1.1430
DATA> 1.1410
DATA> 1.1360
DATA> 1.1395
DATA> 1.1380
DATA> 1.1350
DATA> 1.1370
DATA> 1.1345
DATA> 1.1340
DATA> 1.1405
DATA> 1.1340
DATA> 1.1380
DATA> 1.1355
DATA> end
     16 rows read.
MTB > GStd.
* NOTE  * Standard Graphics are enabled.
          Professional Graphics are disabled.
          Use the GPRO command to enable Professional Graphics.
MTB > hist c1
Character Histogram
Histogram of C1   N = 16
Midpoint   Count
   1.134       3  ***
   1.135       2  **
   1.136       1  *
   1.137       2  **
   1.138       2  **
   1.139       2  **
   1.140       1  *
   1.141       1  *
   1.142       1  *
   1.143       1  *
```

b) Since $\overline{x}$ = 1.1378 and s = 0.0028692 (computed in Minitab), and $t_{0.10/2,15}$ = 1.753, a 90% CI for the mean measured length for bushings of this type is $[1.137, 1.139]$ inches.

c) Since 1.14 in is not in the interval $[1.137, 1.139]$, the supply shop's claim appears to be suspect.

d) Since $\chi^2_{0.01,15}$ = 30.578 and $\chi^2_{0.99,15}$ = 5.229, a 98% CI for σ^2 (the variance of the measured length of the bushings) is

$$\left[\frac{15(0.0028692)^2}{30.578}, \frac{15(0.0028692)^2}{5.229}\right] = \left[4.04 \times 10^{-6}, 23.6 \times 10^{-6}\right]$$

Thus, a 98% CI for the standard deviation is just the square root of the above interval, namely $[0.002, 0.0049]$ inches.

e) In this question, we are interested, basically, in comparing the students' abilities to measure distances. Since the bushing diameters are certainly varying, while we are interested in how the two students compare, the collected data is certainly now paired. Thus, if one wishes to compare the students' average measurements one must look at the *differences* among the individual measurements. Otherwise, one is comparing also the actual differences in bushing diameters.

f) One must assume that the population of differences is normally distributed. When one looks at histogram of the sampled differences (see below), it is difficult to say if this is the case. It is somewhat more normal than found in part (a), but the confidence interval results should still be viewed with a healthy skepticism.

```
NOTE Question #10 Part II
MTB > read c2 c3
DATA> .3690 .3690
DATA> .3690 .3695
DATA> .3690 .3695
DATA> .3700 .3695
DATA> .3695 .3695
DATA> .3700 .3700
DATA> .3695 .3700
DATA> .3690 .3690
DATA> .3690 .3700
DATA> .3695 .3690
DATA> .3690 .3695
DATA> .3690 .3695
DATA> .3695 .3690
DATA> .3700 .3695
DATA> .3690 .3690
DATA> .3690 .3690
DATA> end
     16 rows read.
MTB > let c4=c2-c3
MTB > hist c4
Character Histogram
Histogram of C4   N = 16
Midpoint   Count
 -0.0010       1  *
 -0.0008       0
 -0.0006       4  ****
 -0.0004       1  *
 -0.0002       0
  0.0000       6  ******
  0.0002       0
  0.0004       2  **
  0.0006       2  **
```

g) Minitab's `describe c4` commands (not shown) gives $\bar{d} = -.00093754$ and $s_d = .00045530$. Since $t_{0.025,15} = 2.131$, a 95% CI for the mean difference in outside measurements for the two students is

$$\bar{d} \pm t_{0.025,15}\left(\frac{s}{\sqrt{n}}\right) = [-0.00034, 0.00015] \text{ inches}$$

h) Since 0 is in the interval $[-0.00034, 0.00015]$, we cannot conclude, with 95% confidence, that the students are not coming up with the same mean measurement. (Note the double negative – this is necessitated by the fact that we **can't** conclude that they are coming up with the same mean measurement! The evidence for such a statement would only come after the students had measured an infinite number of samples...

i) We cannot compare the means separately because the observations are not independent. Consider the example where we are trying to compare the measurement skills of two students and ask them to measure a) the width of a hair, b) the height of the CN tower, and c) the distance to the moon. If one compared the averages of these measurement, one would find that the pooled **sample standard deviation** would completely overwhelm the sample mean. Thus, no conclusion could be made. However, considering differences in measurements allows us to restrict our attention to the differences in measurement skills – the standard deviation will be considerably reduced, so that we can make a conclusion at some significance level.

11) a) Let p_H be the probability of obtaining a faulty plastic bag under high-speed operation. A point estimate for p_H is $\hat{p}_H = \frac{147}{250} = 0.588$ and a 95% CI for p_H is $0.588 \pm 1.96\sqrt{\frac{0.588(0.412)}{250}} = [.527, .649]$.

b) Let p_L be the probability of obtaining a faulty plastic bag under low-speed operation. A point estimate for p_L is $\hat{p}_L = \frac{12}{250} = 0.048$ and a 95% CI for p_L is thus $0.048 \pm 1.96\sqrt{\frac{0.048(0.952)}{250}} = [0.022, 0.074]$.

c) A point estimate for $p_H - p_L$ is $\hat{p}_H - \hat{p}_L = 0.588 - 0.048 = 0.54$ and a 95% CI for p_L is $0.54 \pm 1.96\sqrt{\frac{0.588(0.412)}{250} + \frac{0.048(0.952)}{250}} = [0.47, 0.61]$.

d) Since 0 is not in the interval $[0.47, 0.61]$ (and not even close), it seems quite unlikely that the two speeds of operation yield the same probability of obtaining a faulty bag.

12) a) Since $f_{0.05,9,9} = 3.18$ and the sample standard deviations are as calculated in Minitab below, a 90% CI for σ_1^2/σ_2^2, the ratio of the variances from the two hours, is $[\frac{0.00019889^2}{0.00027507^2} \times \frac{1}{3.18}, \frac{0.00019889^2}{0.00027507^2} \times 3.18] = [0.16, 1.66]$ inches^2.

```
MTB > NOTE Question #12
MTB > read 'g:\inwk6211\fittings.dat' c1 c2
Entering data from file: g:\inwk6211\fittings.dat
     10 rows read.
MTB > name c1 'hour1' c2 'hour2'
MTB > stdev c1
Column Standard Deviation
  Standard deviation of hour1   =  0.00019889
MTB > stdev c2
Column Standard Deviation
  Standard deviation of hour2   =  0.00027507
```

b) Based on (a) (since 1 is in the interval $[0.16, 1.66]$), the assumption of equal variances appears reasonable. When one looks at histograms of the sampled data, the assumption of normally distributed populations seems reasonable as well.

```
MTB > hist c1
Character Histogram
Histogram of hour1   N = 10
Midpoint   Count
  0.4032       1  *
  0.4033       2  **
  0.4034       1  *
  0.4035       3  ***
  0.4036       2  **
  0.4037       0
  0.4038       0
  0.4039       1  *

MTB > hist c2
Character Histogram
Histogram of hour2   N = 10
Midpoint   Count
  0.4027       1  *
  0.4028       0
  0.4029       1  *
  0.4030       1  *
  0.4031       2  **
  0.4032       2  **
  0.4033       0
  0.4034       1  *
  0.4035       1  *
  0.4036       1  *
```

c) Since the means and standard deviations are (as computed by Minitab below and above), and $t_{0.02,18} = 2.214$ and

$$s_p^2 = \frac{9(0.00019889)^2 + 9(0.00027507)^2}{10 + 10 - 2} = 5.76 \times 10^{-8}$$

so that $s_p = 0.00024$, a 96% confidence interval for the difference in the two means is

$$(0.40348 - 0.40317) \pm 2.214(0.00024)\sqrt{\frac{1}{10} + \frac{1}{10}} = [0.0000724, 0.000548]$$

```
MTB > mean c1
Column Mean
  Mean of hour1   =     0.40348
MTB > mean c2
Column Mean
  Mean of hour2   =     0.40317
```

d) Since 0 is not in the interval [0.0000724, 0.000548], there appears to be evidence, at the 96% confidence level, that the hourly mean pitch diameter differs.

13) a) Since the sample standard deviation is 0.019599 (as calculated by Minitab below) and $\chi^2_{0.01,15} = 30.578$ and $\chi^2_{0.99,15} = 5.229$, a 98% confidence interval for the variance of the wear is

$$\left[\frac{15(0.019599)^2}{30.578}, \frac{15(0.019599)^2}{5.229}\right] = [0.00019, 0.0011]$$

Thus, a 98% CI for the standard deviation of the wear is [0.014, 0.033].

```
MTB > NOTE Question #13
MTB > read 'g:\inwk6211\valves.dat' c1
Entering data from file: g:\inwk6211\valves.dat
```

```
     16 rows read.
MTB > stdev c1
Column Standard Deviation
   Standard deviation of C1      =    0.019599
```

b) The population of wear in coated valves must be normally distributed in order for our CI's to be accurate. When one looks at a histogram of the sampled data, this assumption seems suspect as the distribution looks skewed. Thus, as usual, we must be skeptical of our results, pending further investigation.

```
MTB > hist c1
Character Histogram
Histogram of C1   N = 16
Midpoint   Count
    0.07       2  **
    0.08       5  *****
    0.09       3  ***
    0.10       3  ***
    0.11       1  *
    0.12       1  *
    0.13       0
    0.14       0
    0.15       1  *
```

14) a) $H_o:\ p = 0.25$ (or $p \geq 0.25$; taking $p = 0.25$ is the worst case)
$H_a:\ p < 0.25$

b) Type I error is rejecting H_o when it is true. For the marketing director, it means concluding that the market share is less than 25%, when in fact it isn't.

c) Type II error is not rejecting H_o when it is false. For the marketing director, it means concluding that the market share is 25%, when in fact it isn't.

d) One could argue both ways. Conceivably, type II error is more serious.

15) a) Let μ_L be the population mean measured length for bushings of this type. Our test is

$H_o:\quad \mu_L = 1.14$ in
$H_a:\quad \mu_L \neq 1.14$ in

at the $\alpha = 0.05$ significance level. Our critical rejection region is $t < -t_{.05/2,16-1} = -t_{.025,15} = -2.131$ and $t > 2.131$. The computed sample statistic (based on the mean and standard deviation calculations by Minitab of the first students measurements, shown below) is

$$t = \frac{1.1378 - 1.14}{0.0028692/\sqrt{16}} = -3.067$$

and thus, since $-3.067 < -2.131$, we reject H_o and conclude, with 95% confidence, that the mean measured length is not 1.14 in. Note that this conclusion is drawn based on the assumption that the lengths of the spanner bushings are normally distributed. Recall in Question #10 that a histogram of the sampled data did not appear to be normally distributed. Thus, our conclusion is suspect. Assuming, however, that the population is normally distributed, the exact p-value (using Minitab's `cdf` command as shown below) is

$$2P[T < -3.067] = 2(0.0039) = 0.0078$$

which says that the evidence is quite definitely in favour of $H_a:\quad \mu_L \neq 1.14$.

```
MTB > read c1
DATA> 1.1375
DATA> 1.1390
DATA> 1.1420
DATA> 1.1430
```

```
DATA> 1.1410
DATA> 1.1360
DATA> 1.1395
DATA> 1.1380
DATA> 1.1350
DATA> 1.1370
DATA> 1.1345
DATA> 1.1340
DATA> 1.1405
DATA> 1.1340
DATA> 1.1380
DATA> 1.1355
DATA> end
     16 rows read.
MTB > mean c1
Column Mean
  Mean of C1      =      1.1378
MTB > stdev c1
Column Standard Deviation
  Standard deviation of C1      =   0.0028692
MTB > cdf -3.067;
SUBC> t 15.
Cumulative Distribution Function
Student's t distribution with 15 d.f.
       x        P(X <= x)
  -3.067          0.0039
```

b) Let μ_1 be the population mean measured length for bushings for Student A and let μ_2 be the population mean measured length for bushings for Student B. We want to test

$$H_o: \quad \mu_1 - \mu_2 = \mu_D = 0$$
$$H_a: \quad \mu_1 - \mu_2 \neq 0$$

at the $\alpha = 0.05$ significance level. Our critical rejection region is $t < -t_{0.05/2,16-1} = -t_{0.025,15} = -2.131$ and $t > 2.131$. The computed sample statistic (based on the mean and standard deviation calculations of Minitab below) is

$$t = \frac{-0.000093754 - 0}{0.00045530/\sqrt{16}} = -0.824$$

and thus, since $-2.131 < -0.824 < 2.131$, we cannot reject H_o with 95% confidence and conclude that there is insufficient evidence to say that the two students' average measurements are not the same. Note that this conclusion is drawn based on the assumption that the difference in the two students' measured lengths are normally distributed. Recall in Question #10 that a histogram of the sampled differences did not appear to be normal. Granted the sample size is small, but the validity of the assumption is suspect so that our conclusion is suspect. However, assuming normality to hold, the exact p-value (using Minitab's cdf as shown below) is

$$2\text{P}\,[T < -0.824] = 2(0.2114) = 0.4228$$

```
MTB > read c2 c3
DATA> .3690 .3690
DATA> .3690 .3695
DATA> .3690 .3695
DATA> .3700 .3695
DATA> .3695 .3695
DATA> .3700 .3700
DATA> .3695 .3700
DATA> .3690 .3690
DATA> .3690 .3700
```

```
DATA> .3695 .3690
DATA> .3690 .3695
DATA> .3690 .3695
DATA> .3695 .3690
DATA> .3700 .3695
DATA> .3690 .3690
DATA> .3690 .3690
DATA> end
     16 rows read.
MTB > let c4=c2-c3
MTB > mean c4
Column Mean
  Mean of C4      =-0.000093754
MTB > stdev c4
Column Standard Deviation
  Standard deviation of C4      =  0.00045530
MTB > cdf -.824;
SUBC> t 15.
Cumulative Distribution Function
Student's t distribution with 15 d.f.
         x       P(X <= x)
   -0.8240        0.2114
```

c) Yes, the same conclusion could have been drawn based on our previously built 95% CI. Since we were testing at α = 0.05 (the same α as in the CI) we merely needed to check to see if the hypothesized value of 0 was included in the CI. Since it is, we cannot reject H_o at α = 0.05.

16) a) Let p_H be the probability of obtaining a faulty bag under high-speed operation. Then, we want to test

$$\begin{aligned} H_o: &\quad p_H = 0.50 \\ H_a: &\quad p_H > .50 \end{aligned}$$

at the α = 0.05 significance level. Our critical rejection region is $z > z_{0.05}$ = 1.645. The computed test statistic is

$$z = \frac{\frac{147}{250} - 0.50}{\sqrt{\frac{0.50(0.50)}{250}}} = 2.78$$

and thus since 2.78 $>$ 1.645 we must reject H_o at the α = 0.05 significance level and conclude that the probability of obtaining a faulty bag under high-speed operation is more than 50%. The p-value is P$[Z > 2.78]$ = 0.0027, which agrees with our conclusion since we reject H_o for any $\alpha >$ p-value.

```
MTB> cdf -2.78;
SUBC> Normal 0 1.
Cumulative Distribution Function
Normal with mean = 0 and standard deviation = 1.00000
         x      P( X <= x)
   -2.7800        0.0027
```

b) Let p_L be the probability of obtaining a faulty bag under low- speed operation. Then, we want to test

$$\begin{aligned} H_o: &\quad p_L = 0.04 \\ H_a: &\quad p_L > 0.04 \end{aligned}$$

at the α = 0.05 significance level. Our critical rejection region is $z > z_{0.05}$ = 1.645. The computed test statistic is

$$z = \frac{\frac{12}{250} - 0.04}{\sqrt{\frac{0.04(0.96)}{250}}} = 0.645$$

and thus, since $0.645 < 1.645$ we cannot reject H_o with 95% confidence and conclude that there is insufficient evidence to show that the probability of obtaining a faulty bag under low-speed operation is more than 4%. The p-value is P $[Z > 0.645] = 0.2595$.

```
MTB > cdf -0.645;
SUBC> Normal 0 1.
Cumulative Distribution Function
Normal with mean = 0 and standard deviation = 1.00000
          x      P( X <= x)
   -0.6450        0.2595
```

c) Now we want to test

$$\begin{aligned} H_o: &\quad p_H = p_L \\ H_a: &\quad p_H > p_L \end{aligned}$$

at the $\alpha = 0.05$ significance level. Our critical rejection region is $z > z_{0.05} = 1.645$. The 'pooled' estimate of the proportion faulty, assuming H_o is true, is

$$\hat{p} = \frac{147 + 12}{250 + 250} = 0.318$$

and our computed test statistic is

$$z = \frac{\frac{147}{250} - \frac{12}{250}}{\sqrt{0.318(0.682)\left(\frac{1}{250} + \frac{1}{250}\right)}} = 18.33$$

and thus, since $18.33 > 1.645$, we must reject H_o at the $\alpha = 0.05$ significance level and conclude that the probabilities of obtaining a faulty bag are not the same at the two different speeds of operation. The p-value is ≈ 0. There is overwhelming evidence that the two speeds yield different proportions of defective plastic bags.

17) Let μ_1 be the population mean pitch diameter for the first hour.
Let μ_2 be the population mean pitch diameter for the second hour.
Then, we want to test

$$\begin{aligned} H_o: &\quad \mu_1 - \mu_2 = 0 \\ H_a: &\quad \mu_1 - \mu_2 \neq 0 \end{aligned}$$

at the $\alpha = 0.05$ significance level. Our critical rejection region is $t < -t_{0.05/2,10+10-2} = -t_{0.025,18} = -2.101$ and $t > 2.101$. The computed test statistic (based on the mean and standard deviation calculations of Minitab below) is

$$t = \frac{.40348 - .40317}{0.00024\sqrt{\frac{1}{10} + \frac{1}{10}}} = 2.888$$

and thus since $2.888 > 2.101$ we must reject H_o at the $\alpha = 0.05$ significance level and conclude that there is an hourly difference in the pitch diameter. The p-value is 2P $[T > 2.888] = 2(0.0049) = 0.0098$.

```
MTB > NOTE Question #17
MTB > read 'g:\inwk6211\fittings.dat' c1 c2
Entering data from file: g:\inwk6211\fittings.dat
     10 rows read.
MTB > name c1 'hour1' c2 'hour2'
MTB > stdev c1
Column Standard Deviation
```

```
  Standard deviation of hour1   =  0.00019889
MTB > stdev c2
Column Standard Deviation
  Standard deviation of hour2   =  0.00027507
MTB > mean c1
Column Mean
  Mean of hour1   =     0.40348
MTB > mean c2
Column Mean
  Mean of hour2   =     0.40317
```