Taylor Maurer, 11/12/2020, ISE201 HW4

8.1.7 *A manufacturer produces piston rings for an automobile engine. It is known that ring diameter is normally distributed with $\sigma = 0.001$ millimeters. A random sample of 15 rings has a mean diameter of $\bar{x} = 74.036$ millimeters.*

1. *Construct a 99% two-sided confidence interval on the mean piston ring diameter.*
   Answer: First we can say α is equal to 1% as 1 - .99 = 0.01. Thus we will be trying to find $z_{.01/2}$. Within the appendix the probability of 0.05 correlates to z value of 2.57. Now we can use the following equation

$$P(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$
$$P(74.036 - (2.57)\frac{0.001}{\sqrt{15}} \leq \mu \leq 74.036 + (2.57)\frac{0.001}{\sqrt{15}}) = 0.99 \tag{1}$$
$$P(74.0353 \leq \mu \leq 74.0366) = 0.99$$

   From the above the lower bound is $L = 74.0353$ and the upper bound is $U = 74.0366$.

2. *Construct a 99% lower-confidence bound on the mean piston ring diameter. Compare the lower bound of this confidence interval with the one in part (a).*
   Answer: First we can say α is equal to 1% as 1 - .99 = 0.01. Thus we will be trying to find $z_{.01}$. Within the appendix the probability of 0.05 correlates to z value of 2.32. Now we can use the following equation

$$P(\bar{x} - z_{\alpha}\frac{\sigma}{\sqrt{n}} \leq \mu) = 1 - \alpha$$
$$P(74.036 - (2.32)\frac{0.001}{\sqrt{15}} \leq \mu) = 0.99 \tag{2}$$
$$P(74.0354 \leq \mu) = 0.99$$

   From the above the lower bound is $L = 74.0354$. This is slightly larger than the lower bound of the two-sided bound problem. Mathematically that is because $z_{.01}$ resulted in a smaller number than that of problem (a). What this means is that since we've taken away the upper bound portion of the 'alpha' space from the normal distribution we were able to put that on the lower bound side. What that means is the lower bound is now closer to the mean, making it larger.

8.2.1 *A random sample has been taken from a normal distribution.*

1. *Fill in the missing quantities.*
   Answer: The variance is simply the standard deviation squared, so $\sigma^2 = 1.605^2 = 2.576025$. The mean is simply the sum divided by N, so $\frac{251.848}{10} = 25.18$.

Variable	N	Mean	SE Mean	StDev	Variance	Sum
x	10	25.18	0.507	1.605	2.576	251.848

2. *Find a 95% CI on the population mean.*

Answer: First we see from the appendix that $x_{\alpha/2} = x_{0.025} = 1.96$. Then within the following equation we substitue in the values:

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$25.18 - 1.96\frac{1.605}{\sqrt{10}} \leq \mu \leq 25.18 + 1.96\frac{1.605}{\sqrt{10}} \tag{3}$$

$$24.185 \leq \mu \leq 26.17$$

8.3.1 *The percentage of titanium in an alloy used in aerospace castings is measured in 51 randomly selected parts. The sample standard deviation is s = 0.37. Construct a 95% two-sided confidence interval for σ.*

Answer: The confidence interval is described below

$$\chi^2|0.025, 50 = 71.42$$

$$\chi^2|1 - 0.025, 50 = 32.36$$

$$CI = \frac{(51-1)(0.37)^2}{71.42} \leq \sigma^2 \leq \frac{(51-1)(0.37)^2}{32.36} \tag{4}$$

$$= 0.0958 \leq \sigma^2 \leq 0.2115$$

8.4.5 *The U.S. Postal Service (USPS) has used optical character recognition (OCR) since the mid-1960s. In 1983, USPS began deploying the technology to major post offices throughout the country (www.britannica.com). Suppose that in a random sample of 500 handwritten zip code digits, 466 were read correctly.*

1. *Construct a 95% confidence interval for the true proportion of correct digits that can be automatically read.*

   Answer: First define phat and z-value.

$$\hat{p} = \frac{X}{n} = \frac{466}{50} = 0.932$$

$$z_{\alpha/2} = 1.96 \tag{5}$$

   Now use the following equation to find the confidence interval:

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.932 - 1.96\sqrt{\frac{0.932(1-0.932)}{500}} \leq p \leq 0.932 + 1.96\sqrt{\frac{0.932(1-0.932)}{500}} \tag{6}$$

$$0.90993 \leq 0.95406$$

2. *What sample size is needed to reduce the margin of error to 1%?.*

   Answer: Error is 0.01. To find the sample size see the following equation:

$$n = (\frac{z_{\alpha/2}}{E})^2 p(1-p)$$

$$n = (\frac{z_{\alpha/2}}{E})^2 \hat{p}(1-\hat{p})$$

$$= (\frac{1.96}{0.01})^2 0.932(1-0.932) \tag{7}$$

$$n = 2434.65 = 2435$$

3. *How would the answer to part (b) change if you had to assume that the machine read only one-half of the digits correctly?*

Answer: $\hat{p}$ would be 0.5 thus:

$$\begin{aligned}
n &= (\frac{z_{\alpha/2}}{E})^2 p(1-p) \\
n &= (\frac{z_{\alpha/2}}{E})^2 \hat{p}(1-\hat{p}) \\
&= (\frac{1.96}{0.01})^2 0.5(1-0.5) \\
n &= 9604
\end{aligned} \tag{8}$$

8.1.10 *If the sample size n is doubled, by how much is the length of the CI on μ in Equation 8.5 reduced? What happens to the length of the interval if the sample size is increased by a factor of four?*

Answer: The length of the interval is reduced by a factor of $\frac{1}{\sqrt{2}}$ if the sample size is doubled. If the sample size is increased by 4, the interval is reduced by factor of $\frac{1}{2}$. See the equation below:

The confidence interval is defined within equation 8.5:

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{9}$$

Now the interval length is defined by the lower bound subtracted by the upper bound:

$$\begin{aligned}
length &= \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} - \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \\
&= -2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}
\end{aligned} \tag{10}$$

If you put in a factor of Cn in for n:

$$\begin{aligned}
&= -2z_{\alpha/2}\frac{\sigma}{\sqrt{Cn}} \\
&= -2z_{\alpha/2}\frac{\sigma}{\sqrt{C}\sqrt{n}} \\
&= \frac{-2}{\sqrt{C}}z_{\alpha/2}\frac{\sigma}{\sqrt{n}}
\end{aligned} \tag{11}$$

Thus when C is equal to 2 the length of the interval is reduced by a factor of $\frac{1}{\sqrt{2}}$ and when C is equal to 4 the length of the interval is reduced by a factor of $\frac{1}{2}$.

8.1.11 *By how much must the sample size n be increased if the length of the CI on μ in Equation 8.5 is to be halved?*

Answer: As shown in the previous problem, 8.1.10, if C becomes 4 the CI is reduced by half. Therefore the answer to the question is the sample size needs to be increased by a factor of 4 for the length of the CI to be reduced by half. Reference solution to 8.1.10.

9.1.3 *A textile fiber manufacturer is investigating a new drapery yarn, which the company claims has a mean thread elongation of 12 kilograms with a standard deviation of 0.5 kilograms. The company wishes to test the hypothesis H0: $\mu = 12$ against H1: $\mu < 12$, using a random sample of four specimens.*

1. *What is the type I error probability if the critical region is defined as $\bar{x} < 11.5$ kilograms?*
   Answer: We first get the following z value:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{11.5 - 12}{0.5/2} = -2 \qquad (12)$$

   This corresponds to an alpha (type-I error probability) of:

$$\alpha = P(Z < -2) = 0.02275 \qquad (13)$$

2. *Find β for the case in which the true mean elongation is 11.25 kilograms.*
   Answer: Beta is defined as the following

$$\beta = P(\bar{x} > 11.5, when \ \mu = 11.25) \qquad (14)$$

   So we can use the following equation, although μ is now μ_{true}

$$z = \frac{\bar{x} - \mu_{true}}{\sigma/\sqrt{n}}$$
$$= \frac{11.5 - 11.25}{.5/2} = 1 \qquad (15)$$

   This cooresponds to:
$$\beta = P(z > 1) = 1 - 0.84135 = 0.15865 \qquad (16)$$

3. *Find β for the case in which the true mean is 11.5 kilograms.*
   Answer: Simply use 11.5 instead of 11.25 in the equation:

$$z = \frac{\bar{x} - \mu_{true}}{\sigma/\sqrt{n}}$$
$$= \frac{11.5 - 11.5}{.5/2} = 0 \qquad (17)$$

   This corresponds to:
$$\beta = P(z > 0) = 1 - 0.5 = 0.5 \qquad (18)$$

9.1.4 *Repeat Exercise 9.1.3 using a sample size of $n = 16$ and the same critical region.*

1. *What is the type I error probability if the critical region is defined as $\bar{x} < 11.5$ kilograms?*
   Answer: We first get the following z value:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{11.5 - 12}{0.5/4} = -4 \qquad (19)$$

   This corresponds to an alpha (type-I error probability) of:

$$\alpha = P(Z < -4) = 0.000033 \qquad (20)$$

2. *Find β for the case in which the true mean elongation is 11.25 kilograms.*
   Answer: Beta is defined as the following

$$\beta = P(\bar{x} > 11.5, when \ \mu = 11.25) \qquad (21)$$

So we can use the following equation, although μ is now μ_{true}

$$
\begin{aligned}
z &= \frac{\bar{x} - \mu_{true}}{\sigma/\sqrt{n}} \\
&= \frac{11.5 - 11.25}{.5/4} = 2
\end{aligned}
\tag{22}
$$

This cooresponds to:

$$
\beta = P(z > 2) = 1 - 0.977250 = 0.02275
\tag{23}
$$

3. *Find β for the case in which the true mean is 11.5 kilograms.*
   Answer: Simply use 11.5 instead of 11.25 in the equation:

$$
\begin{aligned}
z &= \frac{\bar{x} - \mu_{true}}{\sigma/\sqrt{n}} \\
&= \frac{11.5 - 11.5}{.5/4} = 0
\end{aligned}
\tag{24}
$$

This corresponds to:

$$
\beta = P(z > 1) = 1 - 0.5 = 0.5
\tag{25}
$$

9.1.5 *In Exercise 9.1.3, calculate the P-value if the observed statistic is*

1. $\bar{x} = 11.25$
   Answer: We first get the following z value:

$$
z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{11.25 - 12}{0.5/2} = -3
\tag{26}
$$

This corresponds to the following pvalue:

$$
\begin{aligned}
pval &= 1 - P(-3 < Z < 3) \\
&= 1 - (P(Z > -3) + P(Z < 3)) \\
&= 1 - (P(Z < 3) - P(Z < -3)) \\
&= 1 - (0.99865 - 0.001350) \\
&= 0.0027
\end{aligned}
\tag{27}
$$

And for each side of the normal distribution, that would represent 0.00135.

2. $\bar{x} = 11$
   Answer: We first get the following z value:

$$
z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{11 - 12}{0.5/2} = -4
\tag{28}
$$

This corresponds to the following pvalue:

$$
\begin{aligned}
pval &= 1 - P(-4 < Z < 4) \\
&= 1 - (P(Z > -4) + P(Z < 4)) \\
&= 1 - (P(Z < 4) - P(Z < -4)) \\
&= 1 - (0.999952 - 0.000033) \\
&= 0.000081
\end{aligned}
\tag{29}
$$

And for each side of the normal distribution, that would represent 0.0000405.

3. $\bar{x} = 11.75$

Answer: We first get the following z value:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{11.75 - 12}{0.5/2} = -1 \tag{30}$$

This corresponds to the following pvalue:

$$
\begin{aligned}
pval &= 1 - P(-1 < Z < 1) \\
&= 1 - (P(Z > -1) + P(Z < 1)) \\
&= 1 - (P(Z < 1) - P(Z < -1)) \\
&= 1 - (0.841345 - 0.158655) \\
&= 0.31731
\end{aligned}
\tag{31}
$$

And for each side of the normal distribution, that would represent 0.158655.

9.1.6 *In Exercise 9.1.3, calculate the probability of a type II error if the true mean elongation is 11.5 kilograms and*

1. $\alpha = 0.05$ *and* $n = 4$.

    Answer: $\alpha = 0.05$ correspoinds to z value of -1.64. So rearranging the following equation we solve for $\bar{x}$.

    $$
    \begin{aligned}
    z &= \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \\
    \bar{x} &= z\frac{\sigma}{\sqrt{n}} + \mu \\
    &= -1.64\frac{.5}{2} + 12 \\
    \bar{x} &= 11.59
    \end{aligned}
    \tag{32}
    $$

    Now we can find teh corresponding z-value for the β calculation.

    $$
    \begin{aligned}
    z &= \frac{\bar{x} - \mu_{true}}{\sigma/\sqrt{n}} \\
    &= \frac{11.59 - 11.5}{.5/2} = 0.36
    \end{aligned}
    \tag{33}
    $$

    This corresponds to a probability of 0.640576. Now we can find β:

    $$\beta = 1 - P(Z < 0.36) = 1 - 0.640576 = 0.359424 \tag{34}$$

2. $\alpha = 0.05$ *and* $n = 16$.

    Answer: $\alpha = 0.05$ correspoinds to z value of -1.64. So rearranging the following equation we solve for $\bar{x}$.

    $$
    \begin{aligned}
    z &= \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \\
    \bar{x} &= z\frac{\sigma}{\sqrt{n}} + \mu \\
    &= -1.64\frac{.5}{4} + 12 \\
    \bar{x} &= 11.795
    \end{aligned}
    \tag{35}
    $$

Now we can find teh corresponding z-value for the β calculation.

$$z = \frac{\bar{x} - \mu_{true}}{\sigma/\sqrt{n}}$$
$$= \frac{11.795 - 11.5}{.5/4} = 2.36 \tag{36}$$

This corresponds to a probability of 0.990863. Now we can find β:

$$\beta = 1 - P(Z < 0.36) = 1 - 0.990863 = 0.009137 \tag{37}$$

3. *Compare the values of β calculated in the previous parts. What conclusion can you draw?*
   Answer: With an increasing number of samples the value of β decreases. This indicates a the decreasing probability of a type II error and an increasing 'power' of the hypothesis test.

9.5.4 *An article in Fortune (September 21, 1992) claimed that nearly one-half of all engineers continue academic studies beyond the B.S. degree, ultimately receiving either an M.S. or a Ph.D. degree. Data from an article in Engineering Horizons (Spring 1990) indicated that 117 of 484 new engineering graduates were planning graduate study.*

1. *Are the data from Engineering Horizons consistent with the claim reported by Fortune? Use $\alpha = 0.05$ in reaching your conclusions. Find the P-value for this test.*
   Answer: First define the null hypothesis and alternative hypothesis
   $H_0 : p = 1/2$, $H_1 : 0 \neq 1/2$
   Now the estimator gives us:
   $$\hat{p} = 117/484 = 0.2417. \tag{38}$$

   Now we will find the test statistic:

   $$z_0 = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}}$$
   $$= \frac{117 - (484)(1/2)}{\sqrt{(484)(1/2)(1 - 1/2)}} \tag{39}$$
   $$= \frac{-124}{121}$$
   $$= -1.024$$

   Now $|z_0| = 1.024$ corresponds to a probability of 0.8461. This makes our pvalue the following:

   $$P = 2[1 - z_0] = 2[1 - 0.153864] = 0.3078 \tag{40}$$

   This is larger than the α of 0.05, therefore the null hypothesis is not rejected.

2. *Discuss how you could have answered the question in part (a) by constructing a two-sided confidence interval on p.*
   Answer: If you define the confidence interval you will have an upper and lower bound that basically bounds what p could be. If the bounds are more restrictive than the $\alpha = 0.05$ bounds (as in closer to the mean), then you would not be able to reject the null hypothesis even if you did a fixed level significance test.

9.7.2 *Consider the following frequency table of observations on the random variable X.*

1. *Based on these 100 observations, is a Poisson distribution with a mean of 1.2 an appropriate model?*
   Perform a goodness-of-fit procedure with $\alpha = 0.05$.
   Answer: A poison distribution is described via the following:

$$f(x) = \frac{e^{-\lambda T}(\lambda T)^x}{x!} \tag{41}$$

Now the mean is the expected value of λT, the mean is given by:

$$\mu = \frac{(0)(24) + (1)(30) + (2)(31) + (3)(11) + (4)(4)}{24 + 30 + 31 + 11 + 4} = 1.41 \tag{42}$$

So we can estimate the probability for each value by using our estimation parameter within the first equation. I've shown the work for the first two values, and populated the table with the rest:

$$\begin{aligned}
f(x) &= \frac{e^{-\mu}(\mu)^x}{x!} \\
f(0) &= \frac{e^{-1.41}(-1.41)^0}{0!} = 0.244 \\
f(2) &= \frac{e^{-1.41}(-1.41)^0}{1!} = 0.344
\end{aligned} \tag{43}$$

...

Values	Obs. Frequency	Exp. Probability	Exp. Frequency
0	24	0.244	
1	30	0.344	
2	31	0.243	
3	11	0.114	
4	4	0.040	

Now to get the values in the expected frequency column we multiply the expected probability by n, which is 100.

Values	Obs. Frequency	Exp. Probability	Exp. Frequency
0	24	0.244	24.4
1	30	0.344	34.4
2	31	0.243	24.3
3	11	0.114	11.4
4	4	0.040	4

Now we can find the chi-square statistic:

$$\begin{aligned}
\chi_0^2 &= \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \\
&= 0.007 + 0.569 + 1.867 + 0.0144 + 0.0001 \\
&= 2.457
\end{aligned} \tag{44}$$

To understand the goodness of fit we need to have a reference. First though, the degrees of freedom within the problem is given by $k - p - 1 = 5 - 1 - 1 = 3$. We find the $\chi_0^2 = 2.457$ value is bound by $\chi_{0.5,3}^2 = 2.37$ and $\chi_{0.1,3}^2 = 6.25$.

2. *Calculate the P-value for this test.*
   Answer: Based on the previously stated χ^2 answers the P-value for this test is between 0.5 and 0.1. Since that is not less than the 0.05 significance within the problem, we'd reject the null hypothesis.

10.2.11 *The overall distance traveled by a golf ball is tested by hitting the ball with Iron Byron, a mechanical golfer with a swing that is said to emulate the distance hit by the legendary champion Byron Nelson. Ten randomly selected balls of two different brands are tested and the overall distance measured.*

1. *Is there evidence that overall distance is approximately normally distributed? Is an assumption of equal variances justified?*
   Answer: Yes there is evidence that the distribution is normally distributed. The sample variances of the two distributions are found from first finding the mean of the data:

$$\bar{x}_1 = \frac{275 + 286 + 287 + 271 + 283 + 271 + 279 + 275 + 263 + 267}{10}$$
$$\bar{x}_1 = 275.5$$
$$\bar{x}_2 = \frac{258 + 244 + 260 + 265 + 273 + 281 + 271 + 270 + 263 + 268}{10} \tag{45}$$
$$\bar{x}_2 = 265.3$$

Then using the following formula we can find the sample variances:

$$s_1^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1}$$
$$= 64.5$$
$$s_2^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1} \tag{46}$$
$$= 100.9$$

Considering the magnitude of the values within the problem, the sample variances are close enough to assume that the variances are equal between the two populations.

2. *Test the hypothesis that both brands of ball have equal mean overall distance. Use $\alpha = 0.05$. What is the P-value?*
   Answer: First I'll state the null and alternative hypotheses:
   $H_0 : \mu_1 - \mu_2 = 0$ and $H_1 : \mu_1 - \mu_2 \neq 0$
   Next we need to find s_p^2:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
$$= \frac{(10 - 1)64.5 + (10 - 1)100.9}{10 + 10 - 2} \tag{47}$$
$$= 82.7$$
$$s_p = 9.09395$$

Now we can find the T-value associated with the following:

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s_p\sqrt{1/n_1 + 1/n_2}}$$
$$= \frac{275.5 - 265.3 - (0)}{9.09395\sqrt{1/10 + 1/10}} \quad (48)$$
$$= 2.508$$

For $t_0 = 2.508$ and degrees of freedom of 18, this gives a bound of 0.025 and 0.01. Since these values are smaller than 0.05, we'd reject $H0$. For our P-value we take those bounds and double them (double sided test), giving a bound for the P-value of 0.05 and 0.02.

3. *Construct a 95% two-sided CI on the mean difference in overall distance for the two brands of golf balls.*

   Answer: For this problem, our t-value is going to be given by $t_{0.025,18}$. This corresponds to a value of: 2.101. This gets used in the following formula:

$$\bar{x}_1 - \bar{x}_2 - 2.101s_p\sqrt{1/n_1 + 1/n_2} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + 2.101s_p\sqrt{1/n_1 + 1/n_2}$$
$$275.5 - 265.3 - (2.101)(9.09395)\sqrt{1/10 + 1/10} \le \mu_1 - \mu_2 \le 275.5 - 265.3 + (2.101)(9.09395)\sqrt{1/10 + 1/10}$$
$$1.655 \le \mu_1 - \mu_2 \le 18.745$$
$$(49)$$

4. *What is the power of the statistical test in part (b) to detect a true difference in mean overall distance of 5 yards?*

   Answer: First we find d:
$$d = \frac{5}{2s_p} = \frac{5}{2 * 9.09} = 0.275 \quad (50)$$

   Now sample size n^* to use the appropriate curves

$$n^* = 2n - 1 = 19 \quad (51)$$

   Now we use the curves to trace back to β, the closest I can find for the probability of accepting the H0 is 0.95 - 0.9. Then turning that into power we subtract that from one, so the power is approximately 0.05 to 0.1.

5. *What sample size would be required to detect a true difference in mean overall distance of 3 yards with power of approximately 0.75?*

   Answer: Find d again
$$d = \frac{3}{2s_p} = \frac{3}{2 * 9.09} = 0.165 \quad (52)$$

   Now the power is 0.75, which means the probability of accepting H_0 is actually 0.25. When we look at the curves this corresponds to a n^* of 100. Then change this to n via $n = (n^* + 1)/2 = (111)/2 = 55.5$ and then round down to 55.

10.4.2 *A computer scientist is investigating the usefulness of two different design languages in improving programming tasks. Twelve expert programmers who are familiar with both languages are asked to code a standard function in both languages and the time (in minutes) is recorded.*

10

1. *Is the assumption that the difference in coding time is normally distributed reasonable?*

   Answer:
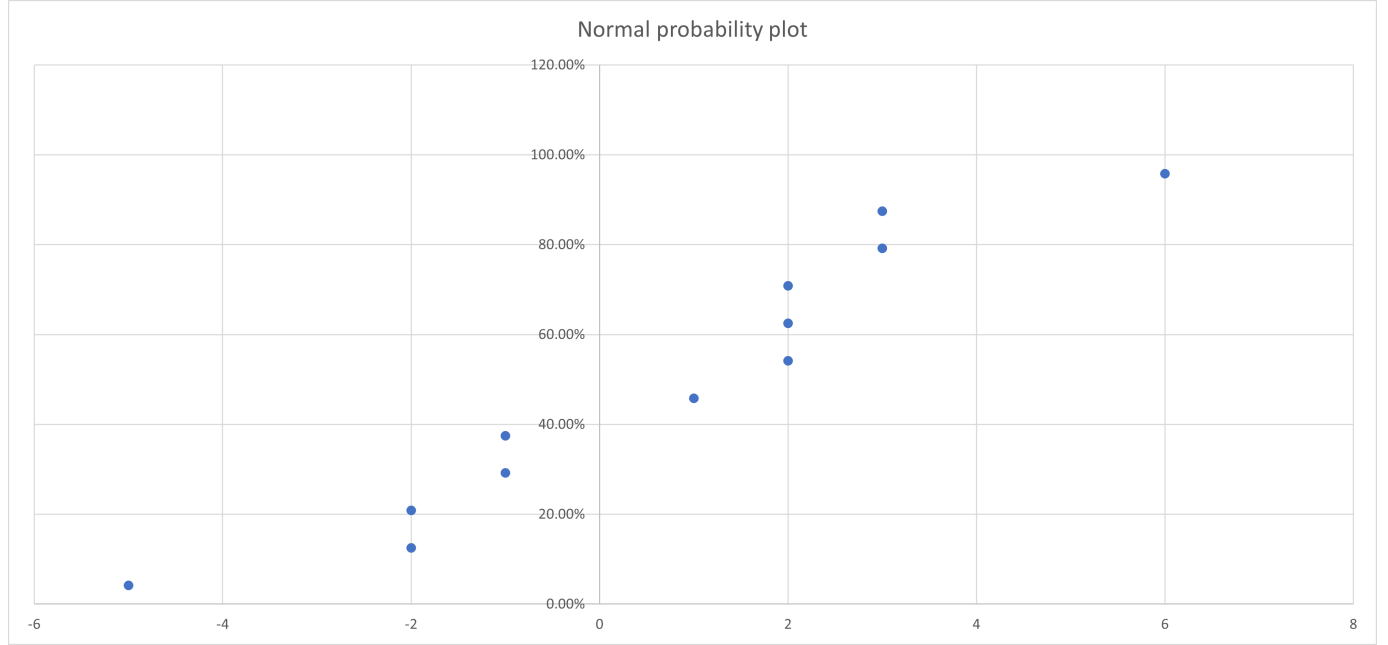
   Shown here is a normal probability plot constructed from the differences of the two data sets. As you can see there's a fairly straight line one could draw proving this distribution is fairly normal. This was generated within Excel.

2. *Find a 95% confidence interval on the difference in mean coding times. Is there any indication that one design language is preferable?*

   Answer: The first thing we can do is find the mean difference $\hat{D}$.

$$
\begin{aligned}
\bar{x}_{DL1} &= \frac{17 + 16 + 21 + 14 + 18 + 24 + 16 + 14 + 21 + 23 + 13 + 18}{12} \\
\bar{x}_{DL1} &= 17.91 \\
\bar{x}_{DL2} &= \frac{18 + 14 + 19 + 11 + 23 + 21 + 10 + 13 + 19 + 24 + 15 + 20}{12} \\
\bar{x}_{DL2} &= 17.25 \\
\bar{D} &= \bar{x}_{DL1} - \bar{x}_{DL2} \\
\bar{D} &= 17.91 - 17.25 = 0.667
\end{aligned}
\tag{53}
$$

Now we find s_D.

$$
\begin{aligned}
s_D^2 &= \frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1} \\
D_i &= X_{i,1} - X_{i,2}... \\
s_D^2 &= 8.78 \\
s_D &= 2.963
\end{aligned}
\tag{54}
$$

Then we want to find the value of $t_{0.025,11}$. This corresponds to a t-value of 2.201. When put within

the following confidence interval equation:

$$\bar{d} - t_{0.025,11}s_d/\sqrt{n} \le \mu_D \le \bar{d} + t_{0.025,11}s_d/\sqrt{n}$$
$$0.667 - (2.201)(8.78)/\sqrt{12} \le \mu_D \le 0.667 + (2.201)(8.78)/\sqrt{12} \tag{55}$$
$$-1.21568 \le \mu_D \le 2.5498$$

There's a little more confidence on the positive side of the normal bell curve. So you could make the argument that as your mean is given by the difference of DL1 - DL2, that the curve skews towards design language 1. But really it's hard to tell from simply the confidence interval alone.

10.4.5 *Neuroscientists conducted research in a Canadian prison to see whether solitary confinement affects brain wave activity ["Changes in EEG Alpha Frequency and Evoked Response Latency During Solitary Confinement," Journal of Abnormal Psychology (1972, Vol. 7, pp. 54–59)]. They randomly assigned 20 inmates to two groups, assigning half to solitary confinement and the other half to regular confinement.*

1. *Is a paired t-test appropriate for testing whether the mean alpha wave frequencies are the same in the two groups? Explain.*
   Answer: The first thing we can do is find the mean difference $\hat{D}$.

$$\bar{x}_{NC} = \frac{10.7 + 10.7 + 10.4 + 10.9 + 10.5 + 10.3 + 9.6 + 11.1 + 11.2 + 10.4}{10}$$
$$\bar{x}_{NC} = 10.58$$
$$\bar{x}_{C} = \frac{9.6 + 10.4 + 9.7 + 10.3 + 9.2 + 9.3 + 9.9 + 9.5 + 9.0 + 10.9}{10} \tag{56}$$
$$\bar{x}_{C} = 9.78$$
$$\bar{D} = \bar{x}_{NC} - \bar{x}_{C}$$
$$\bar{D} = 10.58 - 9.78 = 0.8$$

Then using the mean we can find the variance or standard deviation.

$$s_D^2 = \frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1}$$
$$D_i = X_{i,1} - X_{i,2}... \tag{57}$$
$$s_D^2 = 0.786$$
$$s_D = 0.618$$

This is fairly small variance and standard deviation so maybe the paired t-test isn't the best approach.

2. *Perform an appropriate test.*
   Answer: First I'll state the null and alternative hypotheses:
   $H_0 : \mu_1 - \mu_2 = 0$ and $H_1 : \mu_1 - \mu_2 \ne 0$
   Next we need to use independent t-testing as we already proved the pair testing wouldn't work. So

we need to find s_p. First find s_1 and s_2.

$$s_1^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$
$$s_1^2 = 0.459$$
$$s_1 = 0.211$$
$$s_2^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \tag{58}$$
$$s_2^2 = 0.598$$
$$s_2 = 0.357$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
$$= \frac{(10-1)0.459 + (10-1)0.598}{10 + 10 - 2} \tag{59}$$
$$= 0.5285$$
$$s_p = 0.72698$$

Now we translate that into a t-value.

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s_p \sqrt{1/n_1 + 1/n_2}}$$
$$= \frac{0.8 - (0)}{.0.72698 \sqrt{1/10 + 1/10}} \tag{60}$$
$$= 2.46$$

When looking in the appendix for 18 degrees of freedom (n + n -2), our t-value of 1.9518 falls between 2.101 and 2.552 which correspond to a significance of 0.025 and 0.01. Since we will fall outside of a signifance of 0.05 we should reject our null hypothesis.

10.5.5 *Reconsider the overall distance data for golf balls in Exercise 10.2.11. Is there evidence to support the claim that the standard deviation of overall distance is the same for both brands of balls (use $\alpha = 0.05$)? Explain how this question can be answered with a 95% confidence interval on $\sigma 1/\sigma 2$.*
Answer: First thing to do is state the null and alternate hypothesis:
$H_0 : \sigma_1 = \sigma_2$, $H_1 : \sigma_1 \neq \sigma_2$
Now the formula for actually finding the test statistic F_0 is written (and calculated below). We rely on the two variance values calculated from problem 10.2.11, of $s_1^2 = 64.5$ and $s_2^2 = 100.9$.

$$f_0 = \frac{s_1^2}{s_2^2}$$
$$= \frac{64.5}{100.9} \tag{61}$$
$$= 0.63881$$

Next we need to look up the f-values from the f-distribution. From the appendix (and using 9 degrees of freedom from each independent sample), $f_{0.025,9,9} = 4.03$. Then we use the following equation to find

$f_{1-0.025,9,9} = blank$:

$$f_{1-0.025,9,9} = \frac{1}{f_{0.025,9,9}}$$

$$f_{0.975,9,9} = 0.248139$$

(62)

Now the test statistic calculated $f_0 = 0.63881$ is in between the values of $f_{0.975,9,9} = 0.248139$ and $f_{0.025,9,9} = 4.03$ so the null hypothesis should not be rejected.

To find the 95% confidence interval we just use the following equation

$$\frac{s_1^2}{s_2^2} f_{1-\alpha/2,n_2-1,n_1-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} f_{\alpha/2,n_2-1,n_1-1}$$

$$0.63881(0.248139) \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 0.63881(4.03)$$

(63)

$$0.159 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 2.574$$

10.6.1 *Consider the following computer output.*

1. *Is this a one-sided or a two-sided test?*
   Answer: It is never clearly stated, but I'm assuming that H_0 is defined as $p_1 = p_2$. So therefore the test is a two sided test.

2. *Fill in the missing values.*
   Answer: First missing value is the test statistic Z. Before that we need to find the pooled proportion estimator value, this can be found with the formula below:

   $$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$
   $$= \frac{54 + 60}{250 + 290}$$
   $$= 0.2111$$

   (64)

   Now to find the Z statistic:

   $$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$$
   $$= \frac{.216 - .20897}{\sqrt{.2111(1 - .2111)(1/250 + 1/290)}}$$
   $$= 0.199605$$

   (65)

   Now using $z_0 = 0.199605$ we find that corresponds to a probability of: 0.575345. The p-value is given by $P = 2[1 - z_0] = 2[1 - 0.199605] = 0.84931$.

3. *Can the null hypothesis be rejected?* Assuming the significance is 0.05, the null hypothesis cannot be rejected based on a P-value of 0.84931.

4. *Construct an approximate 90% CI for the difference in the two proportions.* Follow equation below and $z_{\alpha/2} = z_{0.05} = 2.58$.

$$\hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\leq p_1 - p_2 \leq$$

$$\hat{p}_1 - \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \quad (66)$$

$$.216 - .206897 - 2.58\sqrt{\frac{.216(1-.216)}{250} + \frac{.206897(1-.206897)}{290}}$$

$$\leq p_1 - p_2 \leq$$

$$.216 - .206897 + 2.58\sqrt{\frac{.216(1-.216)}{250} + \frac{.206897(1-.206897)}{290}} \quad (67)$$

$$-0.0818655 \leq p_1 - p_2 \leq 0.100072 \quad (68)$$