1. If the cholesterol level of healthy men is normally distributed with a mean of 180 and a standard deviation of 20, and men with cholesterol levels over 225 are diagnosed as not healthy, what is the probability of a type one error?

\[ \alpha = P(\text{Type I error}) = P(\text{Reject } H_0 \text{ when } H_0 \text{ is true}) \\
= P(X > 225 \text{ when } \mu = 180, \sigma = 20) \\
= P(X > 225 \text{ when } X \sim \mathcal{N}(\mu = 180, \sigma = 20)) \\
= P(Z > \frac{225 - 180}{20}) \\
= P(Z > 2.25) = 0.0122 \]

2. If the cholesterol level of healthy men is normally distributed with a mean of 180 and a standard deviation of 20, at what level (in excess of 180) should men be diagnosed as not healthy if you want the probability of a type one error to be 2%?

\[ \alpha = 0.02 = P(X > k \text{ when } X \sim \mathcal{N}(\mu = 180, \sigma = 20)) \\
= P(Z > \frac{k - 180}{20}) \\
\Rightarrow \frac{k - 180}{20} = 2.05 \Rightarrow k = 221 \]

3. If men predisposed to heart disease have a mean cholesterol level of 300 with a standard deviation of 30, above what cholesterol level should you diagnose men as predisposed to heart disease if you want the probability of a type II error to be 1%? (The null hypothesis is that a person is not predisposed to heart disease.)

**Hint:** Men with a cholesterol level over 225 are diagnosed as predisposed to heart disease

\[ H_0: X \leq 225 \]
\[ H_a: X > 225. \]

For what value of \( X \), \( \beta(300) = 0.01 \)

\[ \beta = P(\text{Type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false}) \\
0.01 = P(X < c \text{ when } X \sim \mathcal{N}(\mu = 300, \sigma = 30)) \\
= P(Z < \frac{c - 300}{30}) \\
\Rightarrow -2.33 = \frac{c - 300}{30} \\
\Rightarrow c = 230.1 \approx 230 \]
4. Two inventors have developed a new, energy-efficient lawn mower engine. One inventor says that the engine will run continuously for 5 hours (300 minutes) on a single gallon of regular gasoline. Suppose a random sample of 50 engines is tested. The engines run for an average of 295 minutes, with a standard deviation of 20 minutes. The inventor tests the null hypothesis that the mean run time is 300 minutes against the alternative hypothesis that the mean run time is not 300 minutes, using a 0.05 level of significance. The other inventor says that the new engine will run continuously for only 290 minutes on a gallon of gasoline. Find the power of the test to reject the null hypothesis, if the second inventor is correct.

\[ H_0: \mu = 300 \]
\[ H_a: \mu \neq 300 \]

Given:
\[ \overline{x} = 295 \]
\[ s = 20 \]
\[ n = 50 \]

We want \( 1 - \beta (290) \):

\[ \beta (290) = \Phi \left( z_{\alpha/2} + \frac{\mu_0 - \mu^1}{\sigma \sqrt{n}} \right) - \Phi \left( - z_{\alpha/2} + \frac{\mu_0 - \mu^1}{\sigma \sqrt{n}} \right) = \Phi (5.50) - \Phi (1.58) = 1 - 0.9429 = 0.0571 \]

So, the power of the test is:

\[ 1 - 0.0571 = 0.9429 \]

5. An advertisement for a particular brand of automobile states that it accelerates from 0 to 60 mph in an average of 5.0 seconds. Makers of a competing automobile feel that the true average number of seconds it makes to reach 60 mph from 0 is above 5.0. Suppose the population standard deviation is believed to be 0.43 seconds. We wish to test

\[ H_0: \mu = 5.0 \text{ vs } H_a: \mu > 5.0 \]

How many automobiles are required to satisfy \( \alpha = 0.1 \) and \( \beta (5.5) = 0.01 \)?

\[ Z_\alpha = Z_{0.1} = 1.28 \]
\[ Z_\beta = Z_{0.01} = 2.33 \]
\[ \mu_0 = 5.0 \]
\[ \mu^1 = 5.5 \]

\[ n = \left[ \frac{\sigma (Z_\alpha + Z_\beta)}{\mu_0 - \mu^1} \right]^2 = \left[ \frac{0.43 (1.28 + 2.33)}{5.0 - 5.5} \right]^2 = 9.698 \approx 10 \]

6. A SRS of 500 Connecticut high school students' SAT scores are taken. A teacher believes that the mean will be no more than 450, because that is the mean score for the North Eastern US. If the population standard deviation is 100 and the test rejects the null hypothesis at the 1% level of significance, determine whether this test is sufficiently sensitive (has enough power) to be able to detect an increase of 10 points in the population SAT scores.

\[ H_0: \mu \leq 450 \]
\[ H_a: \mu > 450 \]

Given:
\[ \sigma = 100 \]
\[ n = 500 \]
\[ \alpha = 0.01 \]

\[ \beta (\mu^1) = \Phi \left( z_{\alpha} + \frac{\mu_0 - \mu^1}{\sigma \sqrt{n}} \right) \]
\[ \beta (460) = \Phi \left( 2.33 + \frac{460 - 450}{100 \sqrt{500}} \right) = \Phi (0.09) = 0.5359 \]

So, the power of the test is 0.4641.

The test is not very sensitive to a 10-point increase in mean, since the s.d is very high.
7. A coin is tossed 10 times and $x = 6$ heads are observed. Let $p = P(\text{head})$. Do you believe the coin prefers head with significance level 0.10?

$H_0: p = 0.5$

$H_a: p > 0.5$

Test statistic $= x = \# \text{ of heads}$

Reject $H_0$ if $x \geq c$ for some $c$

$\alpha = P(\text{Type I error})$

$\alpha = P(\text{Reject } H_0 / H_0 \text{ is true})$

$= P(x \geq c \mid x \sim \text{bin}(n=10, p=0.5))$

$= 0.1$

(i.e.) $P(x < c \mid x \sim \text{bin}(n=10, p=0.5)) = 0.9$

$= P(x \leq c-1) = 0.9$

$c-1 = 7 \implies c = 8$

Conclusion: Since $x = 6 < 7$, we fail to reject $H_0$.

8. A major corporation offers a large bonus to all of its employees if at least 80 percent of the corporation's 1,000,000 customers are very satisfied. The company conducts a survey of 100 randomly sampled customers to determine whether or not to pay the bonus. The null hypothesis states that the proportion of very satisfied customers is at least 0.80. If the null hypothesis cannot be rejected, given a significance level of 0.05, the company pays the bonus. Suppose the true proportion of satisfied customers is 0.75. Find the power of the test to reject the null hypothesis.

$\beta(0.75) = 1 - \Phi \left( \frac{0.8 - 0.75 - 1.645}{\sqrt{(0.75)(0.25)/100}} \right)$

$= 1 - \Phi (-0.36)$

$= 0.6406$

Power of the test $= 1 - \beta$

$= 0.3594$

9. Consider the output shown below.

One-Sample Z test
Test of $\mu = 30$ vs $\mu \neq 30$
The assumed standard deviation $\sigma = 1.2$

<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>95% CI</th>
</tr>
</thead>
</table>

(a) Fill in the missing values in the output. What conclusion would you draw?

$Z^* = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{31.2 - 30}{1.2/\sqrt{16}} = 4$

$P\text{-value} = 2 \times P(Z \geq |Z^*|) = 2 \times (1 - \Phi(4)) \approx 0$

(b) Use the output and the normal table to find a 99 percent CI on the mean.

$Z_{0.005} = 2.575$

$(30.4275, 31.9725)$
(c) What is the P-value if the alternative hypothesis is $H_a: \mu > 30$

\[
P\text{-value} = P(Z > Z^*) = P(Z > 4) \approx 0
\]

10. Consider the output shown below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>SE</th>
<th>Mean</th>
<th>95% CI</th>
<th>T*</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>25</td>
<td>92.5805</td>
<td>?</td>
<td>0.4675</td>
<td>(91.6160, ?)</td>
<td>3.38</td>
<td>0.002</td>
<td></td>
</tr>
</tbody>
</table>

(a) Fill in the missing values in the output. Can the null hypothesis be rejected at the 0.05 level? Why?

\[
T^* = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \Rightarrow s = 2.33802
\]

\[
\bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 93.5456
\]

Since $p < 0.05$, reject $H_0$ at 5% level.

(b) Use the output and the $t$ table to find a 99 percent two-sided CI on the mean.

\[
\bar{x} = 0.01
\]

\[t_{0.005, 24} = 2.797\]

\[92.5805 \pm 2.797 \left( \frac{2.33802}{\sqrt{25}} \right)\]

\[(91.2726, 93.8884)\]

11. A police chief claims that the standard deviation in the length of response times is less than 3.7 minutes. A random sample of nine response times has a standard deviation of 3.0 minutes. At $\alpha = 0.05$, is there enough evidence to support the police chief's claim? Assume the population is normally distributed.

$H_0: \sigma = 3.7$

$H_a: \sigma < 3.7$

\[n = 9\]

\[S = 3.0\]

\[\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(8)(3.0)^2}{(3.7)^2} = 5.259\]

\[\chi^2_{0.95, 8} = 2.733\]

Reject $H_0$ if $\chi^2 > \chi^2_{1-\alpha, v}$

Since $\chi^2 > 2.733$, we fail to reject $H_0$. \[v = 8\]
12. You are analyzing a sample of 40 values drawn from a normal distribution with mean 4 and variance 10. \( S^2 \) is the unbiased sample variance. Find the critical value \( c \) such that \( \Pr(S^2 \leq c) = 0.9 \).

\[
\Pr(S^2 \leq c) = 0.9 = \Pr\left( \frac{(n-1)S^2}{\sigma^2} \leq \frac{c(n-1)}{\sigma^2} \right) = 0.9 = \Pr\left( \chi^2 \leq \frac{c(39)}{10} \right) = 0.9
\]

\[
\frac{c(39)}{10} = 50.66 \quad \Rightarrow \quad c = 12.99
\]

13. Two runs of a soft bake are being evaluated. In each run, ten wafers have been baked for 15 minutes at 850°C. Following the soft bake, the resist is etched in an asher. The observed resist etch rates (in Angstroms/sec) are shown below in Table below.

<table>
<thead>
<tr>
<th>Etch Rate After Soft Bake Run 1 (Angstroms/sec)</th>
<th>Etch Rate After Soft Bake Run 2 (Angstroms/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>389.8</td>
<td>441.5</td>
</tr>
<tr>
<td>387.9</td>
<td>331.5</td>
</tr>
<tr>
<td>379.3</td>
<td>335</td>
</tr>
<tr>
<td>354</td>
<td>461.0</td>
</tr>
<tr>
<td>395</td>
<td>313</td>
</tr>
<tr>
<td>289</td>
<td>370</td>
</tr>
<tr>
<td>392</td>
<td>395.8</td>
</tr>
<tr>
<td>295.5</td>
<td>408.5</td>
</tr>
<tr>
<td>369.5</td>
<td>451</td>
</tr>
<tr>
<td>341</td>
<td>394.5</td>
</tr>
</tbody>
</table>

1. Calculate the mean and the variance of the etch rates of each run.

\[
\bar{X}_1 = 359.3 \quad \quad \bar{X}_2 = 390.18
\]

\[
S^2_1 = 1552.39 \quad \quad S^2_2 = 2731.81
\]

2. Perform an F-test to investigate if the two variances are equal.

\[
H_0 : \sigma^2_1 = \sigma^2_2
\]

\[
H_a : \sigma^2_1 \neq \sigma^2_2
\]

\[
F^* = \frac{S^2_1}{S^2_2} = \frac{1552.39}{2731.81} = 0.5683
\]

\[
(\text{or}) \quad F^* \leq F_{1-\alpha/2, m-1, n-1} = 3.18
\]

We fail to reject \( H_0 \).
Since \(6_1^2 = 6_2^2\) in (2), we need to find the pooled variance
\[ s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = 2.1421 \]
\[ d.f = n_1+n_2-2 \]

3. Based on the observation can we claim that both runs of the soft bake have the same mean etch rate? (Using \(\alpha=0.05\))

\[ H_0: \mu_1 = \mu_2 \]
\[ H_a: \mu_1 \neq \mu_2 \]

\[ T^* = \frac{359.3 - 390.18}{\sqrt{2142.1 \left( \frac{1}{10} + \frac{1}{10} \right)}} = -1.492 \]
\[ t_{0.025,18} = 2.101 \]

Reject \(H_0\) if \(T^* \geq t_{d/2, d.f}\) or \(T^* \leq -t_{d/2, d.f}\)

4. Find the 95\% confidence interval on the mean difference etch rate.

\[ \alpha = 0.05 \]
\[ t_{0.025,18} = 2.101 \]

\[ 359.3 - 390.18 \pm 2.101 \sqrt{2142.103 \left( \frac{1}{10} + \frac{1}{10} \right)} \]
\[ (-74.367, 12.607) \]

14. If you walk toward a squirrel that is on the ground, it will eventually run for the nearest tree. A researcher wondered if he could get closer to the squirrel then the squirrel was to the nearest tree before the squirrel would run. He made 11 observations. The data are summarized below:

<table>
<thead>
<tr>
<th>Squirrel</th>
<th>From Person (Y_1)</th>
<th>From Tree (Y_2)</th>
<th>Difference (D=Y_1-Y_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>81</td>
<td>137</td>
<td>-56</td>
</tr>
<tr>
<td>2</td>
<td>178</td>
<td>34</td>
<td>144</td>
</tr>
<tr>
<td>3</td>
<td>202</td>
<td>51</td>
<td>151</td>
</tr>
<tr>
<td>4</td>
<td>325</td>
<td>50</td>
<td>275</td>
</tr>
<tr>
<td>5</td>
<td>238</td>
<td>54</td>
<td>184</td>
</tr>
<tr>
<td>6</td>
<td>134</td>
<td>236</td>
<td>-102</td>
</tr>
<tr>
<td>7</td>
<td>240</td>
<td>45</td>
<td>195</td>
</tr>
<tr>
<td>8</td>
<td>326</td>
<td>293</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>60</td>
<td>277</td>
<td>-217</td>
</tr>
<tr>
<td>10</td>
<td>119</td>
<td>83</td>
<td>36</td>
</tr>
<tr>
<td>11</td>
<td>189</td>
<td>41</td>
<td>148</td>
</tr>
<tr>
<td>Mean</td>
<td>190</td>
<td>118</td>
<td>72</td>
</tr>
</tbody>
</table>

\[ \bar{d} = 72 \]
\[ S_d = 148.014 \]
\[ n = 11 \]
\[ t_{0.025,10} = 2.228 \]

\[ 72 \pm 2.228 \left( \frac{148.014}{\sqrt{11}} \right) \]
\[ (-27.431, 171.431) \]

Construct and interpret a 95\% confidence interval for the mean difference in distance from person to squirrel and squirrel to tree.