Math 326 – Homework 06 (10.2 – 10.8)
Due (via upload to Canvas) Wednesday, March 9, 2022 at 11:59 PM

1. Suppose that we want to study whether students enrolled in a campus wellness program get more sleep than the average college student. (This problem assumes that the wellness group comes from the same general population as the other college students.) Suppose that the average student gets 6.1 hours of sleep per day and let μ denote the average sleep gotten by a student in the wellness population.

   (a) State the hypothesis test for the test you are designing.

   Solution The hypothesis test would be as follows:

$$H_0 : \mu = 6.1$$
$$H_a : \mu > 6.1$$

   Assume the standard deviation is known as $\sigma = 0.5$ hour and that the underlying sleep distribution is normal and that there are 16 students in the wellness program. We define the rejection region to be $\bar{Y} > 6.3$.

   (b) Find α, the probability of Type I error.

   Solution As the underlying distribution is normally distributed, so is $\bar{Y}$ for any sample size. Thus, the rejection region of $\bar{Y} > 6.3$ is equivalent to a z-score of

$$\frac{6.3 - 6.1}{\frac{.5}{\sqrt{16}}} = \frac{.8}{.5} = 1.6 > Z.$$

   Then, we can use a table to see that $P(Z > 1.6) \approx 0.0548 = \alpha$.

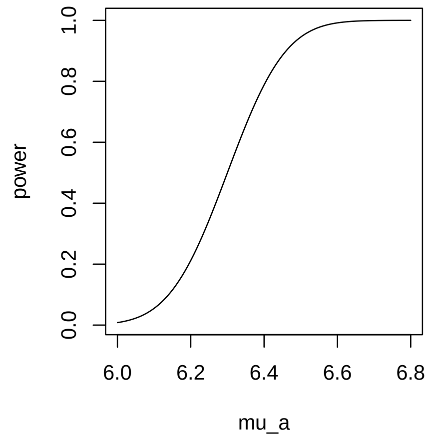
   (c) Find β, the probability of Type II error at $\mu = 6.3, 6.5$, and 6.7.

   Solution If $\mu = 6.3$, then by the symmetry of the normal distribution,
   $P(\bar{Y} < 6.3 | \ \mu = 6.3) = 0.5 = \beta$.
   If $\mu = 6.5$, $P(\bar{Y} < 6.3 | \ \mu = 6.5) = P(Z < -1.6) \approx 0.0548 = \beta$.
   If $\mu = 6.7$, $P(\bar{Y} < 6.3 | \ \mu = 6.7) = P(Z < -3.2) \approx 0.000687 = \beta$.

   (d) Sketch the power function for this test.

   Solution Note this is a one-sided test. So, via R-commands:

```
eq=function(x){pnorm(6.3,x,0.125,lower.tail = FALSE)}
curve(eq,from=6.0, to=6.8,xlab="mua",ylab="power")
```
   we get



2. Lifetimes of a certain brand of switches follows (approximately) a normal distribution with mean 100 hours. Five switches of a new brand are obtained and tested and their lifetimes are measured to be 120, 101, 114, 95, and 130 hours. Does this provide strong evidence that the new switches have a longer average lifespan?

   Solution While the switches are assumed to be near-normal, with such a small sample size we need to use t-statistics. For the given sample, $\bar{X} = 112$ while $s_x = 14.1598$. Because there are 5 data points, a T distribution with 4 degrees of freedom will be used. The T statistic will be $t = \dfrac{112 - 100}{\frac{14.1598}{\sqrt{5}}} \approx 1.895$. Using R, we compute that $P(T > 1.895) \approx 0.065$. This means that if the new switches have the same average lifespan, there would be about a 6.5% chance that a sample of five switches would have this high of a mean or higher. The question of if this is strong evidence that the switches have a longer lifespan is a subjective one, but since $0.05 < 0.065$ I would not say that this provides strong evidence that the new switches have a longer lifespan.

3. A federal regulatory agency hypothesises that the average length of a stay in the hospital is in excess of 5 days. A pilot data set had standard deviation of 3.1 days. Using this as the population standard deviation, how large a sample would you need in designing a test with $\alpha = 0.01$ and $\beta = 0.05$ if the true average is 5.5 days.

   Solution For this test, the hypothesis test will be as follows:

$$H_0 : \mu = 5$$
$$H_a : \mu > 5$$

   Then, if the true average is denoted by $\mu_* = 5.5$, we will look for $\alpha = 0.01 = P(Z > z_1)$

with $z_1 = \dfrac{k-5}{\frac{3.1}{\sqrt{n}}}$ where k is the critical value for the critical region of size 0.01 for $N\left(5, \dfrac{3.1^2}{n}\right)$ and $\beta = 0.05 = P(Z < z_2)$ with $z_2 = \dfrac{k-5.5}{\frac{3.1}{\sqrt{n}}}$. In this second equation for β, our k is the critical value for the critical region of size .05 for $N\left(5.5, \dfrac{3.1^2}{n}\right)$ which will also be the same as our k in the equation for α. To get $\alpha = 0.01$ and $\beta = 0.05$, we also need to have $z_1 = 2.326$ and $z_2 = -1.645$ as $P(Z > 2.326) = 0.01$ and $P(Z < -1.645) = 0.05$. This then gives us

$$2.326 = \frac{k-5}{\frac{3.1}{\sqrt{n}}}$$
$$-1.645 = \frac{k-5.5}{\frac{3.1}{\sqrt{n}}}$$

Solving these equations for k we see that

$$k = \frac{3.1 \cdot 2.326}{\sqrt{n}} + 5$$
$$k = \frac{3.1 \cdot -1.645}{\sqrt{n}} + 5.5$$

Setting these equations equal to each other and solving for n we see that

$$\begin{aligned} \frac{3.1 \cdot 2.326}{\sqrt{n}} + 5 &= \frac{3.1 \cdot -1.645}{\sqrt{n}} + 5.5 \\ \frac{3.1 \cdot 3.971}{\sqrt{n}} &= .5 \\ \sqrt{n} &= \frac{3.1 \cdot 3.971}{.5} \\ n = \left(\frac{3.1 \cdot 3.971}{.5}\right)^2 &\approx 606.15 \end{aligned}$$

So, taking $n = 607$ people will give us a test with $\alpha = 0.01$ and $\beta = 0.05$ if the true mean is 5.5.

4. A study was performed to compare two cholesterol-reducing drugs. Observations of the number of units of cholesterol reduction were recorded for 12 subjects receiving Drug A and 14 subjects receiving Drug B:

	Drug A	Drug B
n	12	14
Mean	5.64	5.03
stnd deviation	1.25	1.82

Researchers are interested in testing if the drugs appear to be *different* in their average cholesterol reduction.

(a) State the assumptions needed for the independent samples t test to be valid.

Solution The samples need to be independent, taken from the same population, and the population needs to be nearly normally distributed.

(b) Perform the t-test, find the p-value, and state the conclusion using a 5% significance level.

Solution For this test, the hypothesis test will be:

$$H_0 : \mu_A - \mu_B = 0$$
$$H_a : \mu_A - \mu_B \neq 0$$

and the test statistic will be $t = \dfrac{5.64 - 5.03}{s_p\sqrt{\dfrac{1}{12} + \dfrac{1}{14}}}$ with $12 + 14 - 2 = 24$ degrees of freedom. In this case, $S_p = \sqrt{\dfrac{11 \cdot 1.25^2 + 13 \cdot 1.82^2}{24}} \approx 1.58441$, so using a calculator we see $t \approx .9787$. Then, technology can be used to see that $P(|t| > .9787) \approx 0.3375$. This means the p-value is $p = 0.3375$ and the conclusion is that there is not enough evidence to reject the null hypothesis at the 5% significance level.