AP Statistics

Name: ______________________________

Type I & II Errors and Power

1. Normal body temperature varies according to a normal distribution with $\mu = 98.6^\circ$ and $\sigma = 0.8^\circ$. A researcher suspects that the mean body temperature of world class runners is higher than that of the general population. He collects a sample of 25 world class runners and finds the mean $\bar{x}$. How large would $\bar{x}$ need to be in order to reject the null hypothesis ($H_O : \mu = 98.6^\circ$) at the $\alpha = 5\%$ level? Find the "rejection region" that is the interval containing $\bar{x}$ such that we would reject H_O. Sketch a curve representing the distribution of $\bar{x}$ along with the "rejection region".

2. The truth of the null hypothesis is unknown and we decide to either reject it or fail to reject it depending on the sample we draw and the magnitude of the p-value. The table below gives all combinations of decisions and truth regarding H_O.

<table>
<tr><td></td><td></td><th colspan="2">Truth about H_O</th></tr>
<tr><td></td><td></td><td>H_O true</td><td>H_O false</td></tr>
<tr><th rowspan="2">Decision</th><td>Reject H_O</td><td></td><td></td></tr>
<tr><td>Fail to reject H_O</td><td></td><td></td></tr>
</table>

Two of the possible decisions are correct and the other two are decision errors.

3. Type I Error

A Type I error occurs when the null hypothesis is true and the decision is to reject it. In our example, what is the probability of making a Type I error? That is, rejecting the null hypothesis when it is, in fact, true?

In general: The probability of committing a Type I error is. . . _________

4. Type II Error

A Type II error occurs when the null hypothesis is false and we fail to reject it. In our example, what is the probability of failing to reject the null hypothesis when it is, in fact, false? What does it depend on? Draw a diagram that shows the area of a Type II error when the true mean body temperature is 99˚. Be sure to keep the same "rejection line" as you had in problem 1.

5. Power of a significance Test

The power of a test is the probability of correctly rejecting the null hypothesis when it is false. Label this area on your diagram above.

6. The mean life of Sylvania light bulbs is advertised to be μ = 500 hours. A consumers' group suspects that the advertisers claim is too high that in fact the mean life span is less than 500 hours. They take a simple random sample of 12 light bulbs and test them. They light each of the 12 bulbs and let them burn until they burn out. They record how long each bulb remained lit until burning out.

Here are the data (hours until burn out).

455	472	472	478	492	493
496	496	509	510	511	525

a. State the null and alternative hypotheses.

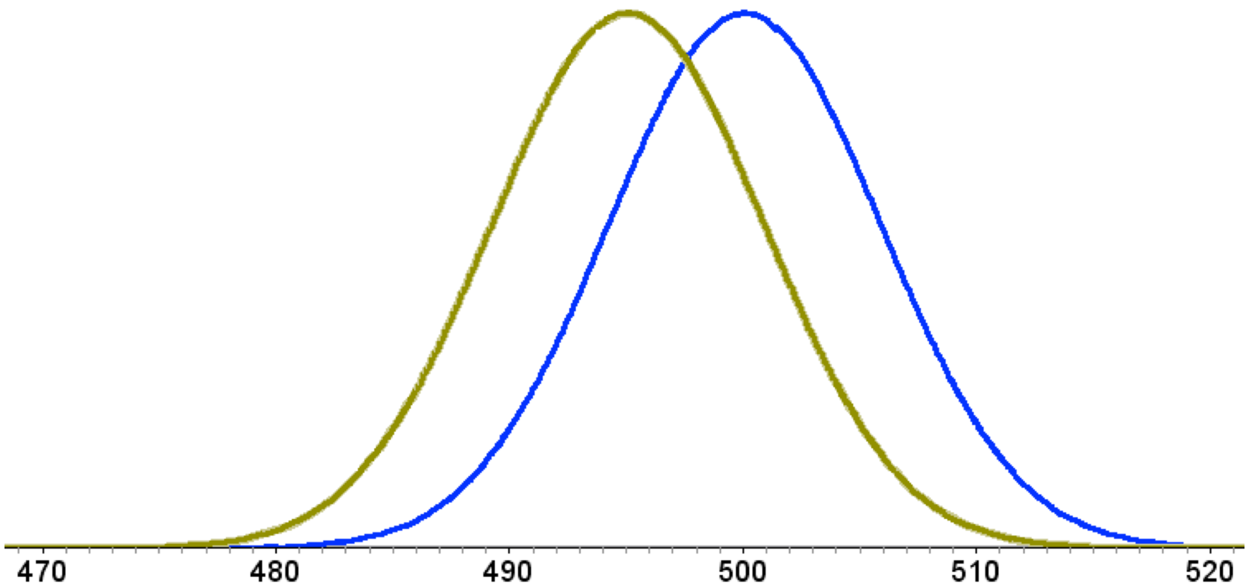
b. Check the conditions necessary to ensure that *t* procedures can be used.

c. Find the rejection region for a significance test at the 10% level.

d. Calculate the sample mean $\bar{x}$. Is it in the rejection region?

e. What conclusion do we draw?

f. Suppose that the null hypothesis is false. That is, $\mu < 500$. Suppose that, in reality, $\mu = 495$. Find the probability of committing a Type II error and the Power of the test. Use and label the diagram below. ***(FYI: On the AP Exam, students will not have to find the probability of a Type II error from "scratch." They should, however, be able to compute the probability of a Type II error given Power, and vice versa.)***

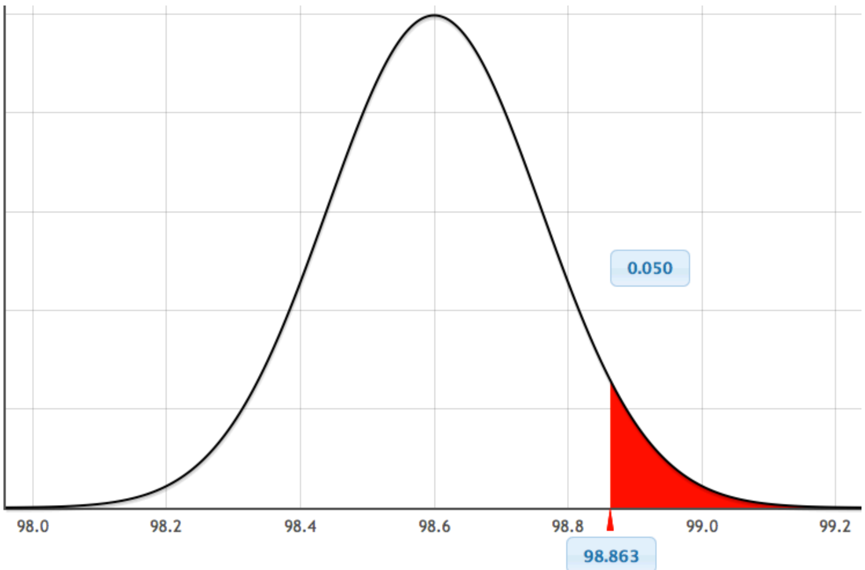


g. Now, suppose that $\mu = 491$ hrs. The probability of making a Type II error is 0.41. What is the power of this test? What is the probability of making a Type I error?

h. You should have discovered that in part g. above, the power increased from the power in part f. Why did this happen? What can be done to increase power even more?

ANSWERS to Type I and Type II Errors and Power:

1. The sampling distribution for x-bar is below: N(98.6, .8/√25). The z-score for a 5% right tail is z = 1.645, and *x-bar* = 98.863˚. The region to the right of 98.863˚ will be the 5% rejection region.

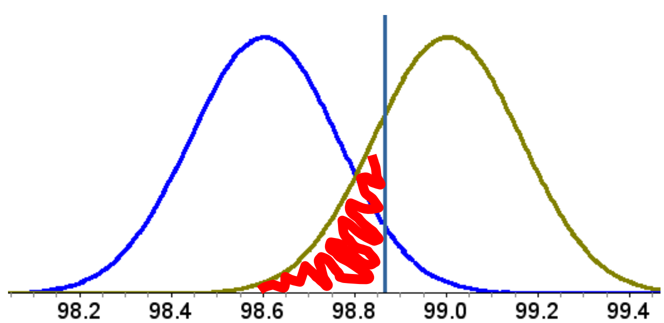


2.

	Truth about H_O	
Decision	H_O true	H_O false
Reject H_O	**Type I Error**	Correct Decision
Fail to reject H_O	Correct Decision	**Type II Error**

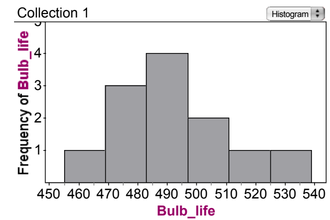
3. The picture drawn IS a picture of when the null hypothesis is true. Therefore, the probability of rejecting the true null hypothesis is the area of the "rejection region," which is 5%. This also happens to be the significance level, alpha! So the probability of a Type I Error is always alpha.

4. Since a Type II error is the probability of not rejecting a ***false*** null hypothesis about the true mean, the probability of a Type II error cannot be calculated unless we know the ***true*** mean (99˚). Once we know the true mean, we can draw a diagram showing the null hypothesis model (blue) vs. the true mean model (olive). The probability of a Type II error will be the area under the true model, but to the left of the rejection line (≈ 20%, shaded)



5. The power is the area to the right of the rejection line, but under the true mean model ≈ 80%.

6. a) $H_0 : \mu = 500 \quad H_A : \mu < 500$, where μ is the mean hours that Sylvania life bulbs last.



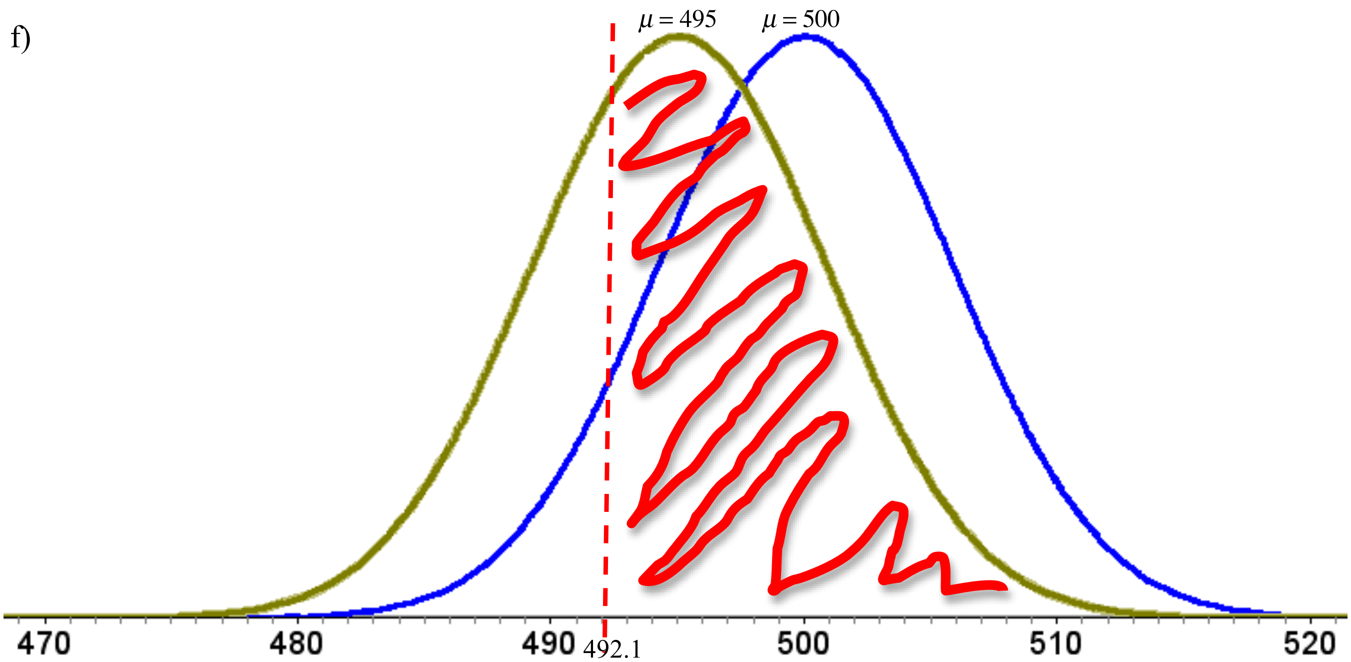
b) Conditions: The sample was stated to be random. The sample size is not large, but a histogram of the sample shows no heavy skewness or outliers, so we can proceed with a t-test. Surely 12 light bulbs is less than 10% of the populations of Sylvania light bulbs.

c) Since $t_{11} = -1.3634 = \frac{\bar{x} - 500}{\frac{20.1}{\sqrt{12}}}$, the rejection region for alpha = 10% is $\bar{x} \leq 492.1$ hrs.

d) NO. $\bar{x} \approx 492.4$ is greater than 492.1, so it is NOT in the rejection region.

e) We cannot reject the null hypothesis since our sample mean of 492.4 is NOT in the rejection region. Also, the p-value is ≈ 0.11, which is higher than the alpha level of 0.10.

f)



P(Type II error) = the area to the right of the rejection line, but under the true mean model (shaded).
$P(t_{11} > \frac{492.1 - 495}{\frac{20.1}{\sqrt{12}}}) = 0.686$. Therefore, power = 1 – 0.686 = 0.314.

g) Power = 1 – P(Type II error) = 1 – 0.41 = 0.59. The probability of making a Type I error is still 10% since alpha is still 10%.

h) When the true mean is farther away from the hypothesized mean, it is easier to detect (i.e. more likely reject) a false null hypothesis. Another way of saying it is that when your standard for rejection is farther away from the truth, you will tend to reject a false null more often. Another way to increase power is to increase your sample size. This will make both distributions more narrow, which will result in higher power (keep in mind the "rejection line" still keeps the same alpha level). This can be demonstrated better with numerous online applets.