STAT 410

Fall 2020
A. Stepanov

Examples for 11/09/2020 (1)

	H_0 true	**H_0 is NOT true**
Do Not Reject H_0	☺	Type II Error
Reject H_0	Type I Error	☺

$$\alpha = \text{significance level} = P(\text{Type I Error}) = P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

$$\beta = P(\text{Type II Error}) = P(\text{Do Not Reject } H_0 \mid H_0 \text{ is NOT true})$$

$$\text{Power} = 1 - P(\text{Type II Error}) = P(\text{Reject } H_0 \mid H_0 \text{ is NOT true})$$

1. A car manufacturer claims that, when driven at a speed of 50 miles per hour on a highway, the mileage of a certain model follows a normal distribution with mean $\mu_0 = 30$ miles per gallon and standard deviation $\sigma = 4$ miles per gallon. A consumer advocate thinks that the manufacturer is overestimating average mileage. The advocate decides to test the null hypothesis $H_0: \mu = 30$ against the alternative hypothesis $H_1: \mu < 30$.

0a) Suppose the actual overall average mileage μ is indeed 30 miles per gallon. What is the probability that the sample mean is 29.4 miles per gallon or less, for a random sample of $n = 25$ cars?

$$P(\overline{X} \le 29.4) = P\left(Z \le \frac{29.4 - 30}{4/\sqrt{25}}\right) = P(Z \le -0.75) = \Phi(-0.75) = \mathbf{0.2266}.$$

0b) A random sample of 25 cars yields $\overline{x}$ = 29.4 miles per gallon. Based on the answer for part (a), is there a reason to believe that the actual overall average mileage is not 30 miles per gallon?

If $\mu = 30$, it is not unusual to see the values of the sample mean $\overline{x}$ at 29.4 miles per gallon or even lower. It does not imply that $\mu = 30$, but we have no reason to doubt the manufacturer's claim.

0c) Suppose the actual overall average mileage μ is indeed 30 miles per gallon. What is the probability that the sample mean is 28 miles per gallon or less, for a random sample of $n = 25$ cars?

$$P(\overline{X} \le 28) = P\left(Z \le \frac{28-30}{4/\sqrt{25}}\right) = P(Z \le -2.50) = \Phi(-2.50) = \mathbf{0.0062}.$$

0d) A random sample of 25 cars yields $\overline{x}$ = 28 miles per gallon. Based on the answer for part (c), is there a reason to believe that the actual overall average mileage is not 30 miles per gallon?

If $\mu = 30$, it is very unusual to see the values of the sample mean $\overline{x}$ at 28 miles per gallon or lower. It does not imply that $\mu < 30$, but we have a very good reason to doubt the manufacturer's claim.

a) Suppose the consumer advocate tests a sample of $n = 25$ cars. What is the significance level associated with the rejection region "Reject H_0 if $\overline{x} < 28.6$"?

α = significance level = P(Type I Error) = P(Reject $H_0 \mid H_0$ true).

Need $P(\overline{X} < 28.6 \mid \mu = 30) = ?$ $\qquad \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} = Z.$

$$P(\overline{X} < 28.6 \mid \mu = 30) = P\left(Z < \frac{28.6-30}{4/\sqrt{25}}\right) = P(Z < -1.75) = \Phi(-1.75) = \mathbf{0.0401}.$$

b) Suppose the consumer advocate tests a sample of $n = 25$ cars. Find the rejection region with the significance level $\alpha = 0.05$.

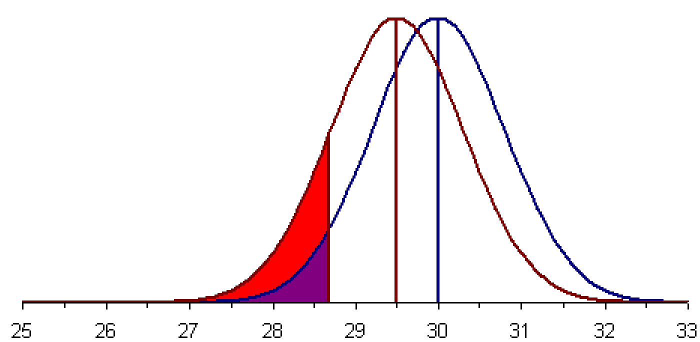
$n = 25.$ $\alpha = 0.05.$

Rejection Region:

Reject H_0 if $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha.$ $Z = \dfrac{\overline{X} - 30}{4 / \sqrt{25}} < -1.645.$

$$\overline{X} < 30 - 1.645 \cdot \frac{4}{\sqrt{25}} = 28.684.$$

c) Suppose the consumer advocate tests a sample of $n = 25$ cars and uses a 5% level of significance. Find the power of the test if the true mean is $\mu_1 = 29.5$.



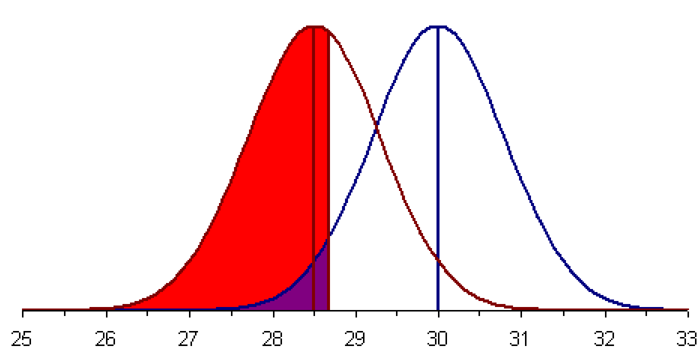
$$P(\overline{X} < 28.684 \mid \mu_1 = 29.5)$$

$$= P\left(Z < \frac{28.684 - 29.5}{4 / \sqrt{25}}\right)$$

$$= P(Z < -1.02)$$

$= \mathbf{0.1539}.$

d) Repeat part (c) for the case when the true value of the mean is $\mu_1 = 28.5$.



$$P(\overline{X} < 28.684 \mid \mu_1 = 28.5)$$

$$= P\left(Z < \frac{28.684 - 28.5}{4 / \sqrt{25}}\right)$$

$$= P(Z < 0.23)$$

$= \mathbf{0.5910}.$

e) Repeat parts (b) – (d) using a 10% level of significance.

$n = 25.$ $\alpha = 0.10.$

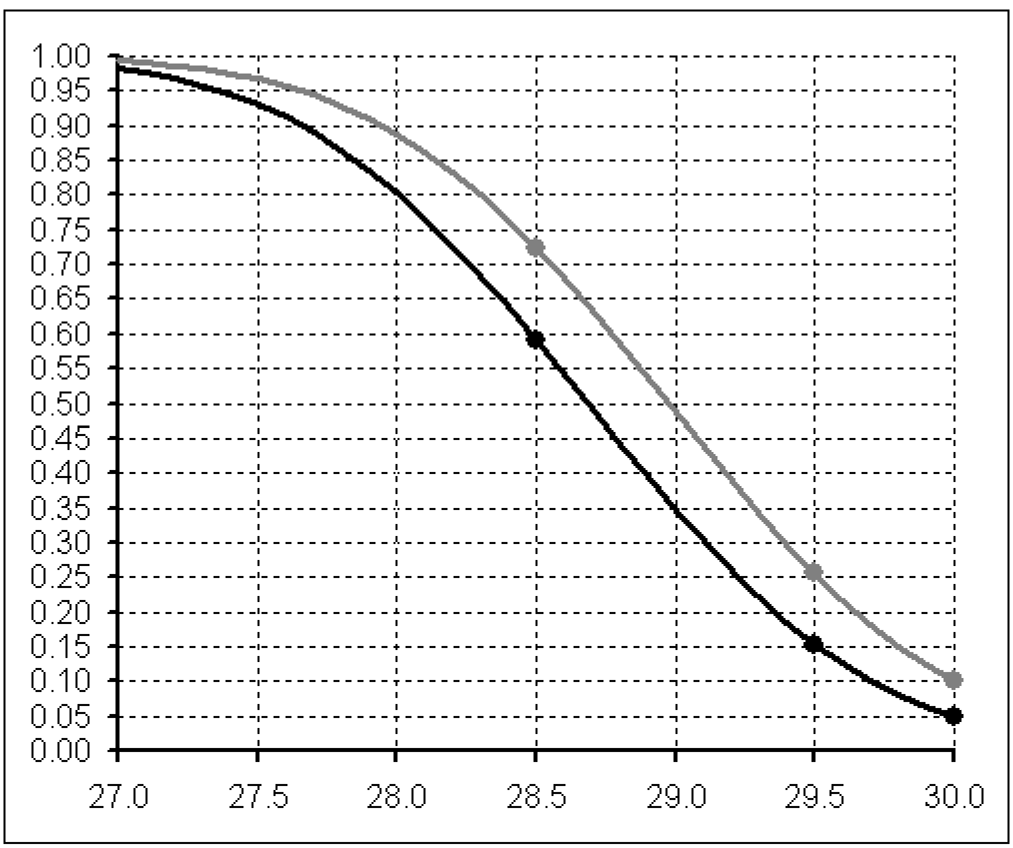
Rejection Region:

Reject H_0 if $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha.$ $Z = \dfrac{\overline{X} - 30}{4 / \sqrt{25}} < -1.282.$

$$\overline{X} < 30 - 1.282 \cdot \frac{4}{\sqrt{25}} = 28.9744.$$

$$P(\overline{X} < 28.9744 \mid \mu_1 = 29.5) = P\left(Z < \frac{28.9744 - 29.5}{4 / \sqrt{25}} \right) = P(Z < -0.657) \approx \mathbf{0.2546}.$$

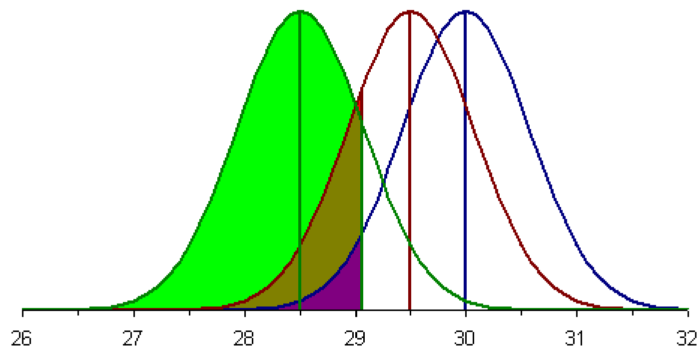
$$P(\overline{X} < 28.9744 \mid \mu_1 = 28.5) = P\left(Z < \frac{28.9744 - 28.5}{4 / \sqrt{25}} \right) = P(Z < 0.593) \approx \mathbf{0.7224}.$$



f) Repeat parts (b) – (d) using a larger sample size of $n = 49$.

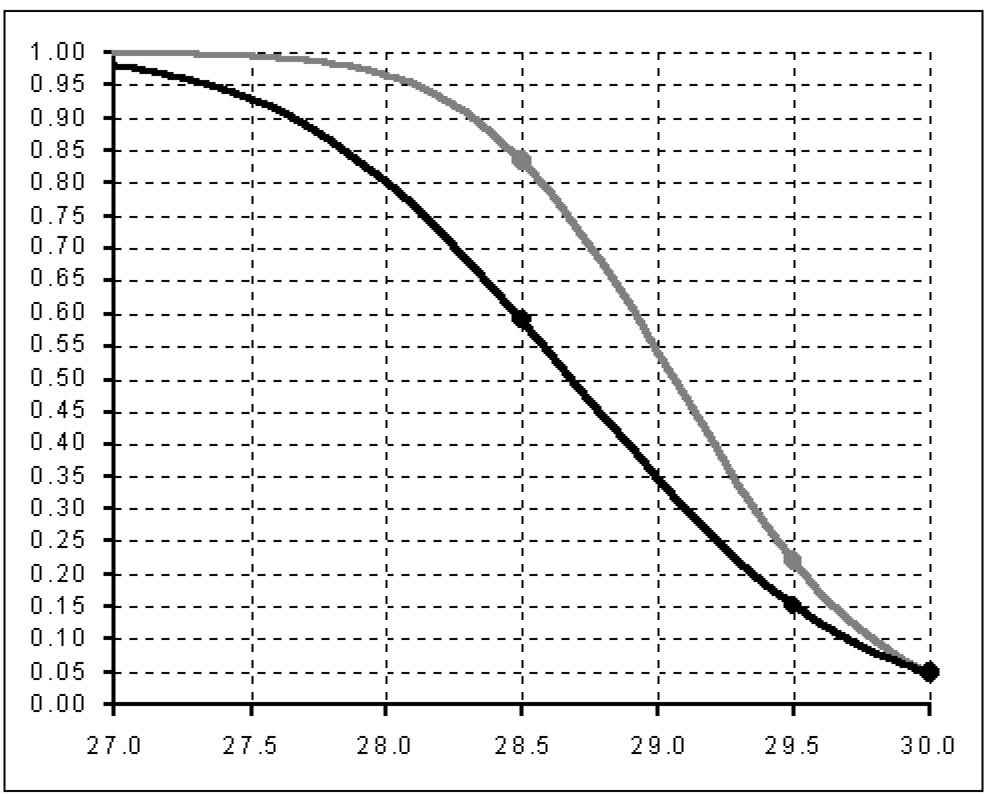
Rejection Region:

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha. \qquad Z = \frac{\overline{X} - 30}{4 / \sqrt{49}} < -1.645. \qquad \overline{X} < 30 - 1.645 \cdot \frac{4}{\sqrt{49}} = 29.06.$$



$$P(\overline{X} < 29.06 \mid \mu_1 = 29.5) = P\left(Z < \frac{29.06 - 29.5}{4 / \sqrt{49}} \right) = P(Z < -0.77) = \mathbf{0.2206}.$$

$$P(\overline{X} < 29.06 \mid \mu_1 = 28.5) = P\left(Z < \frac{29.06 - 28.5}{4 / \sqrt{49}} \right) = P(Z < 0.98) = \mathbf{0.8365}.$$



g) What is the minimum sample size required if we want to have the power of at least 0.80 at $\mu_1 = 29.5$ for the test with a 5% level of significance?

Rejection Region:

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha. \qquad \overline{X} < 30 - 1.645 \cdot \frac{4}{\sqrt{n}}.$$

Want $P\left(\overline{X} < 30 - 1.645 \cdot \frac{4}{\sqrt{n}} \;\middle|\; \mu = 29.5 \right) \ge 0.80.$

$$P\left(\overline{X} < 30 - 1.645 \cdot \frac{4}{\sqrt{n}} \;\middle|\; \mu = 29.5 \right) = P\left(Z < \frac{30 - 29.5}{4/\sqrt{n}} - 1.645 \right)$$

$$= P\left(Z < \frac{\sqrt{n}}{8} - 1.645 \right)$$

Then, since $P(Z < 0.84) \approx 0.80$, $\frac{\sqrt{n}}{8} - 1.645 \ge 0.84.$

$\Rightarrow \quad \sqrt{n} \ge 8 \cdot (0.84 + 1.645) = 19.88.$

$\Rightarrow \quad n \ge 19.88^2 = 395.2144.$ Round up. $n \ge \mathbf{396}.$

1. (continued)

Suppose that the sample mean is $\overline{x} = 29$ miles per gallon for a sample of $n = 25$ cars.

h) Find the p-value of the appropriate test.

$H_0 : \mu \ge 30$ vs. $H_1 : \mu < 30.$ Left – tailed.

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{29 - 30}{4/\sqrt{25}} = \mathbf{-1.25}.$$

$$P(Z \le -1.25) = \Phi(-1.25) = \mathbf{0.1056}.$$

i) State your decision $(\text{Accept } H_0 \text{ or Reject } H_0)$ for the significance level $\alpha = 0.05$.

P-value > α $\Rightarrow$ Accept H_0 P-value < α $\Rightarrow$ Reject H_0

Since $0.1056 > 0.05$, **Do NOT Reject H_0 at $\alpha = 0.05$**.

j) Construct a 95% confidence interval for the overall average miles-per-gallon rating for this model, μ.

$\sigma = 4$ is known. $n = 25$. The confidence interval : $\overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$.

95% confidence level, $\alpha = 0.05$, $\alpha/2 = 0.025$, $z_{\alpha/2} = 1.96$.

$29 \pm 1.96 \cdot \frac{4}{\sqrt{25}}$ **29 ± 1.568** **(27.432 ; 30.568)**

k) What is the minimum sample size required if we want to estimate μ to within 0.5 miles per gallon with 95% confidence?

$\varepsilon = 0.5$, $\sigma = 4$,

95% confidence level, $\alpha = 0.05$, $\alpha/2 = 0.025$, $z_{\alpha/2} = 1.96$.

$n = \left[\frac{z_{\alpha/2} \cdot \sigma}{\varepsilon}\right]^2 = \left[\frac{1.96 \cdot 4}{0.5}\right]^2 = \mathbf{245.8624}$. Round <u>up</u>. $n = \mathbf{246}$.

l) Construct a 95% confidence upper bound for μ.

The confidence upper bound for μ : $\overline{X} + z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}$.

95% confidence level, $\alpha = 0.05$, $z_{\alpha} = 1.645$.

$29 + 1.645 \cdot \frac{4}{\sqrt{25}}$ **29 + 1.316** **(0 ; 30.316)**

Note that $\mu_0 = 30$ is covered. Recall part (i).