# Review Key

1. Suppose the amount spent on rent (in dollars) for NCSU students per month is normally distributed with a mean of \$535 and a standard deviation of \$75.
   a. What is the probability of randomly selecting an NCSU student whose monthly rent is more than \$600?

$$\boldsymbol{X \sim N(535, 75)}$$

$$\boldsymbol{P(X > 600) = P\left(Z > \frac{600 - 535}{75}\right) = P(Z > 0.87) = 1 - P(Z < 0.87)}$$

$$\boldsymbol{= 1 - 0.8078 = 0.1922}$$

   b. If a random sample of 100 NCSU students is collected, what is the probability that the mean monthly rent for these students is less than \$515?

$$\boldsymbol{\bar{X} \sim N(535, \tfrac{75}{\sqrt{100}})} \textbf{ since X is Normal}$$

$$\boldsymbol{P(\bar{X} < 515) = P\left(Z < \frac{515 - 535}{\frac{75}{\sqrt{100}}}\right) = P(Z < -2.67) = 0.0038}$$

2. Megacorp is interested in hiring NC State students about what matters to them more when being hired, Salary or Work Culture. They believe the true proportion of students who will say that salary matters more is 90%. A senior executive conducts a random sample of 200 students, and finds that only 84% of students surveyed said salary was more important. Does this seem reasonable due to random variability alone?

**Belief:** $\boldsymbol{p > 0.90}$

**Observed:** $\boldsymbol{n = 200, \hat{p} = 0.84}$

**What does "reasonable" mean? Statistically, we might think of it instead as: Is it likely to observe a sample proportion of 0.84 (or something more extreme) in a sample of 200, if the true proportion is 0.90 (or higher)?**

**Sampling distribution of the sample proportion is Normal if conditions np >10 and nq>10. Since this is the case here, we can find the probability using the Normal distribution as follows:**

$$\boldsymbol{P(\hat{p} < 0.84) = P\left(Z < \frac{0.84 - 0.90}{\sqrt{\frac{0.90 * 0.10}{200}}}\right) = P(Z < -2.83) = 0.0023}$$

**The probability of observing 84% or less in a sample of 200 is very small, which leads us to believe that it is *not* due to random chance alone.**

3. The p-value in a hypothesis test
   a. is calculated assuming the null hypothesis is false
   b. is the probability that the null hypothesis is correct
   c. is not important for a hypothesis test, only for confidence intervals
   **d. can be used to make a conclusion about our hypotheses.**

4. A farmer is interested in the proportion of corn in his corn fields that are stricken with a certain disease. He randomly selects 90 ears of corn and sees if they have the disease. It turns out 15 of them do. Construct and interpret a 99% confidence interval for the true proportion of corn that has the disease.

$$\hat{p} = \frac{15}{90} \approx 0.1667$$

$$n = 90 \qquad c = 0.99 \qquad z_{.995} = 2.575$$

$$\hat{p} \pm z_c \sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.1667 \pm 2.575 \cdot \sqrt{\frac{(0.1667)(0.8333)}{90}} = 0.1667 \pm 0.1012$$

$$99\%\ CI\ for\ p: (0.0655, 0.2679)$$

**Interpretation:**

**We are 99% confident that the true proportion of corn that has the disease is between 0.0655 and 0.2679.**

5. Grand Auto Corporation produces auto batteries. The company claims that its top-of-the-line batteries are good, on average, for at least 65 months. A consumer protection agency tested 45 such batteries to check this claim. It found that the mean life of these 45 batteries is 63.4 months, and the standard deviation is 3 months. Test the auto corporation's claim at the $\alpha$ = 0.05 level.

$$H_0: \mu \geq 65$$
$$H_a: \mu < 65$$
$$\alpha = 0.05$$

$$df = n - 1 = 44$$

$$Test\ statistic: t^* = (\bar{x} - \mu)/(s/\sqrt{n}) = (63.4 - 65)/(3/\sqrt{45}) = -3.5777$$

$$P - Value: P(t_{44} < -3.578) \approx P(t_{40} < -3.6) = 0.000$$

$$Decision: Since\ p - value < \alpha, we\ will\ reject\ H_0\ in\ favor\ of\ H_a.$$

**Interpretation: There is enough evidence at the α = 0.05 level to suggest that the mean life of these auto batteries is less than 65 months.**

6. Direct Mailing Company sells computers and computer parts by mail. The company claims that at least 90% of all orders are mailed within 72 hours after they are received. The quality control department at the company often takes samples to check if this claim is valid. A recent sample of 150 orders showed that 129 of them were mailed within 72 hours. Do you think the company's claim is true? Use a 1% significance level.

$$H_0: p \geq .90$$

$$H_a: p < .90$$

$$\alpha = 0.01$$

$$Test\ statistic: z^* = \frac{\hat{p} - p}{\sqrt{\frac{p * q}{n}}} = \frac{\frac{129}{150} - 0.90}{\sqrt{\frac{0.90 * 0.10}{150}}} = -1.633$$

$$P - value: P(Z < -1.63) = 0.0516$$

$$Decision: Since\ p - value > \alpha, we\ will\ fail\ to\ reject\ H_0.$$

**Interpret: There is not enough evidence at the 0.01 level to reject the company's claim that the proportion of computer orders that are mailed with 72 hours is at least 0.90.**

7. A biology class is asked to find the average wingspan of monarch butterflies. The class caught and measured the wingspans of 24 monarch butterflies. Their mean wingspan is found to be 93.5 mm with a standard deviation of 3.44 mm. Find a 95% confidence interval for the mean wingspan of all monarch butterflies.

$$n = 24$$

$$\bar{x} = 93.5$$

$$s = 3.44$$

$$df = n - 1 = 23$$

$$c = .95$$

$$t_{0.975,\ 23} = 2.069$$

$$\bar{x} \pm t_c * \frac{s}{\sqrt{n}} = 93.5 \pm 2.069 * \frac{3.44}{\sqrt{24}}$$

$$= 93.5 \pm 1.4526 = (92.0474, 94.9526)$$

**With 95% confidence, we can say that the true population mean of the wingspan of monarch butterflies is between 92.0474 mm and 94.9526 mm.**

8. **True** or False: P-value is the lowest significance level α that results in rejecting the null hypothesis.

9. **True** or False: if the P-value is small, it means that the observed data is very unlikely to have occurred under $H_0$, so we are more confident in rejecting the null hypothesis.

10. **True** or False: a 95% confidence level means that if we were to take 100 different samples and compute a 95% confidence interval for each sample, then approximately 95 of the 100 confidence intervals will contain the true parameter value.