Test Review

Sampling Distributions

1-Prop-Z-Int

1-Prop-Z-Test
1. We have calculated a 95% confidence interval and would prefer for our next confidence interval to have a smaller margin of error without losing any confidence. In order to do this, we can

I. change the z* value to a smaller number.
II. take a larger sample.
III. take a smaller sample.

A) I only  B) II only  C) III only  D) I and II  E) I and III
2. Which is true about a 98% confidence interval for a population proportion based on a given sample?

I. We are 98% confident that other sample proportions will be in our interval.

II. There is a 98% chance that our interval contains the population proportion.

III. The interval is wider than a 95% confidence interval would be.

A) None    B) I only    C) II only    D) III only    E) I and II
3. We have calculated a confidence interval based on a sample of size \( n = 100 \). Now we want to get a better estimate with a margin of error that is only one-fourth as large. How large does our new sample need to be?

A) 25  
B) 50  
C) 200  
D) 400  
E) 1600
4. A certain population is bimodal. We want to estimate its mean, so we will collect a sample. Which should be true if we use a large sample rather than a small one?

I. The distribution of our sample data will be more clearly bimodal.

II. The sampling distribution of the sample means will be approximately normal.

III. The variability of the sample means will be smaller.

A) I only
B) II only
C) III only
D) II and III
E) I, II, and III
5. The manager of an orchard expects about 70% of his apples to exceed the weight requirement for “Grade A” designation. At least how many apples must he sample to be 90% confident of estimating the true proportion within ±4%?

A) 19  B) 23  C) 89  D) 356  E) 505
6. A *P*-value indicates

A) the probability that the null hypothesis is true.

B) the probability that the alternative hypothesis is true.

C) the probability the null is true given the observed statistic.

D) the probability of the observed statistic given that the null hypothesis is true.

E) the probability of the observed statistic given that the alternative hypothesis is true.
7. A statistics professor wants to see if more than 80% of her students enjoyed taking her class. At the end of the term, she takes a random sample of students from her large class and asks, in an anonymous survey, if the students enjoyed taking her class. Which set of hypotheses should she test?

A) $H_0 : p < 0.80$  $H_A : p > 0.80$

B) $H_0 : p = 0.80$  $H_A : p > 0.80$

C) $H_0 : p > 0.80$  $H_A : p = 0.80$

D) $H_0 : p < 0.80$  $H_A : p \neq 0.80$

E) $H_0 : p = 0.80$  $H_A : p < 0.80$
The President’s job approval rating is always a hot topic. Your local paper conducts a poll of 100 randomly selected adults to determine the President’s job approval rating. A CNN/USA Today/Gallup poll conducts a poll of 1010 randomly selected adults.

Which poll is more likely to report that the President’s approval rating is below 50%, assuming that his actual approval rating is 54%? Explain.
• The smaller poll would have more variability and would thus be more likely to vary from the actual approval rating of 54%. We would expect the larger poll to be more consistent with the 54% rating. So, it is more likely that the smaller poll would report that the President’s approval rating is below 50%.
A box of Raspberry Crunch cereal contains a mean of 13 ounces with a standard deviation of 0.5 ounce. The distribution of the contents of cereal boxes is approximately Normal.

What is the probability that a case of 12 cereal boxes contains a total of more than 160 ounces?
**Method 1:**
Let $B =$ weight of one box of cereal and $T =$ weight of 12 boxes of cereal.
We are told that the contents of the boxes are approximately Normal, and we can assume that the content amounts are independent from box to box.

\[ E(T) = E(B_1 + B_2 + \cdots + B_{12}) = E(B_1) + E(B_2) + \cdots + E(B_{12}) = 156 \text{ ounces} \]

Since the content amounts are independent,

\[ Var(T) = Var(B_1 + B_2 + \cdots + B_{12}) = Var(B_1) + Var(B_2) + \cdots + Var(B_{12}) = 3 \]

\[ SD(T) = \sqrt{Var(T)} = \sqrt{3} = 1.73 \text{ ounces} \]

We model $T$ with $N(156, 1.73)$.

\[ z = \frac{160 - 156}{1.73} = 2.31 \text{ and } P(T > 160) = P(z > 2.31) = 0.0104 \]

There is a 1.04\% chance that a case of 12 cereal boxes will weigh more than 160 ounces.
Method 2:
Using the Central Limit Theorem approach, let $\bar{y}$ = average content of boxes in the case. Since the contents are Normally distributed, $\bar{y}$ is modeled by $N\left(13, \frac{0.5}{\sqrt{12}}\right)$.

$$P\left(\bar{y} > \frac{160}{12}\right) = P\left(\bar{y} > 13.33\right) = P\left(z > \frac{13.33 - 13}{0.5/\sqrt{12}}\right) = P\left(z > 2.31\right) = 0.0104.$$  

There is a 1.04% chance that a case of 12 cereal boxes will weigh more than 160 ounces.
A recent Gallup poll found that 28% of U.S. teens aged 13-17 have a computer with Internet access in their rooms. The poll was based on a random sample of 1028 teens and reported a margin of error of ±3%.

What level of confidence did Gallup use for this poll?

\[ ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}, \]  

we have \[ z^* = \frac{0.03}{\sqrt{\frac{0.28 \times 0.72}{1028}}} \approx 2.14. \]

Our confidence level is approximately \( P(-2.14 < z < 2.14) = 0.9676, \) or 97%. 
1. We have calculated a confidence interval based upon a sample of \( n = 200 \). Now we want to get a better estimate with a margin of error only one fifth as large. We need a new sample with \( n \) at least...

A) 40    B) 240    C) 450    D) 1000    E) 5000
2. A certain population is strongly skewed to the right. We want to estimate its mean, so we will collect a sample. Which should be true if we use a large sample rather than a small one?

I. The distribution of our sample data will be closer to normal.
II. The sampling model of the sample means will be closer to normal.
III. The variability of the sample means will be greater.

A) I only  (B) II only  C) III only  D) I and III only  E) II and III only
4. Which is true about a 95% confidence interval based on a given sample?

I. The interval contains 95% of the population.
II. Results from 95% of all samples will lie in the interval.
III. The interval is narrower than a 98% confidence interval would be.

A) None          B) I only     C) II only     D) III only     E) II and III only
5. A truck company wants on-time delivery for 98% of the parts they order from a metal manufacturing plant. They have been ordering from Hudson Manufacturing but will switch to a new, cheaper manufacturer (Steel-R-Us) unless there is evidence that this new manufacturer cannot meet the 98% on-time goal. As a test the truck company purchases a random sample of metal parts from Steel-R-Us, and then determines if these parts were delivered on-time. Which hypothesis should they test?

A) $H_0 : p < 0.98 \quad H_A : p > 0.98$
B) $H_0 : p > 0.98 \quad H_A : p = 0.98$
C) $H_0 : p = 0.98 \quad H_A : p < 0.98$
D) $H_0 : p = 0.98 \quad H_A : p \neq 0.98$
E) $H_0 : p = 0.98 \quad H_A : p > 0.98$
7. A pharmaceutical company investigating whether drug stores are less likely than food markets to remove over-the-counter drugs from the shelves when the drugs are past the expiration date found a $P$-value of 2.8%. This means that:

A) 2.8% more drug stores remove over-the-counter drugs from the shelves when the drugs are past the expiration date.

B) 97.2% more drug stores remove over-the-counter drugs from the shelves when the drugs are past the expiration date than drug stores.

C) There is a 2.8% chance the drug stores remove more expired over-the-counter drugs.

D) There is a 97.2% chance the drug stores remove more expired over-the-counter drugs.

E) None of these.
8. To plan the course offerings for the next year a university department dean needs to estimate what impact the “No Child Left Behind” legislation might have on the teacher credentialing program. Historically, 40% of this university’s pre-service teachers have qualified for paid internship positions each year. The Dean of Education looks at a random sample of internship applications to see what proportion indicate the applicant has achieved the content-mastery that is required for the internship. Based on these data he creates a 90% confidence interval of (33%, 41%). Could this confidence interval be used to test the hypothesis $H_0: p = 0.40$ versus $H_A: p < 0.40$ at the $\alpha = 0.05$ level of significance?

A) Yes, since 40% is in the confidence interval he accepts the null hypothesis, concluding that the percentage of applicants qualified for paid internship positions will stay the same.

B) Yes, since 40% is in the confidence interval he fails to reject the null hypothesis, concluding that there is not strong enough evidence of any change in the percent of qualified applicants.

C) Yes, since 40% is not the center of the confidence interval he rejects the null hypothesis, concluding that the percentage of qualified applicants will decrease.

D) No, because he should have used a 95% confidence interval.

E) No, because the dean only reviewed a sample of the applicants instead of all of them.
A state has two aquariums that have dolphins, with more births recorded at the larger aquarium than at the smaller one. Records indicate that in general babies are equally likely to be male or female, but the gender ratio varies from season to season.

Which aquarium is more likely to report a season when over two-thirds of the dolphins born were males? Explain.

• The smaller aquarium would experience more variability in the season percentage of male births. We would expect the larger aquarium to stay more consistent and closer to the 50-50 ratio for gender births.
A newspaper article reported that a poll based on a sample of 1150 residents of a state showed that the state’s Governor’s job approval rating stood at 58%. They claimed a margin of error of ±3%.

What level of confidence were the pollsters using?

Approval rating: Since $ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$, we have $0.03 = z^* \sqrt{\frac{(0.58)(0.42)}{1150}}$ or $z^* \approx 2.06$.

Confidence level is 96%
A can of pumpkin pie mix contains a mean of 30 ounces and a standard deviation of 2 ounces. The contents of the cans are normally distributed.

What is the probability that four randomly selected cans of pumpkin pie mix contain a total of more than 126 ounces?
Method 1:
Let \( P = \) one can of pumpkin pie mix and \( T = \) four cans of pumpkin pie mix.
We are told that the contents of the cans are normally distributed, and can assume that the content amounts are independent from can to can.
\[
E(T) = E(P_1 + P_2 + P_3 + P_4) = E(P_1) + E(P_2) + E(P_3) + E(P_4) = 120 \text{ ounces}
\]
Since the content amounts are independent,
\[
Var(T) = Var(P_1 + P_2 + P_3 + P_4) = Var(P_1) + Var(P_2) + Var(P_3) + Var(P_4) = 16
\]
\[
SD(T) = \sqrt{Var(T)} = \sqrt{16} = 4 \text{ ounces}
\]
We model \( T \) with \( N(120, 4) \)
\[
z = \frac{126 - 120}{4} = 1.5 \quad P(T > 126) = P(z > 1.5) = 0.067
\]
There is a 6.7% chance that four randomly selected cans of pumpkin pie mix contain more than 126 ounces.
Method 2:
Using the Central Limit Theorem approach, let $\bar{y} =$ average content of cans in sample

Since the contents are Normally distributed, $\bar{y}$ is modeled by $N(30, \frac{2}{\sqrt{4}})$.

$$P\left(\bar{y} > \frac{126}{4}\right) = P(\bar{y} > 31.5) = P\left(z > \frac{31.5 - 30}{1}\right) = P(z > 1.5) = 0.067$$

There is about a 6.7% chance that 4 randomly selected cans will contain a total of over 126 ounces.
A recent psychiatric study from the University of Southampton observed a higher incidence of depression among women whose birth weight was less than 6.6 pounds than in women whose birth weight was over 6.6 pounds. Based on a P-value of 0.0248 the researchers concluded there was evidence that low birth weights may be a risk factor for susceptibility to depression.

Explain in context what the reported P-value means.

• If birth weight was not a risk factor for susceptibility to depression, an observed difference in incidence of depression this large (or larger) would occur in only 2.48% of such samples.
On many highways state police officers conduct inspections of driving logbooks from large trucks to see if the trucker has driven too many hours in a day. At one truck inspection station they issued citations to 49 of 348 truckers that they reviewed.

a. Based on the results of this inspection station, construct and interpret a 95% confidence interval for the proportion of truck drivers that have driven too many hours in a day.

b. Explain the meaning of “95% confidence” in part A.
Conditions:
* Independence: We assume that one trucker’s driving times do not influence other trucker’s driving times.
* Random Condition: We assume that trucks are stopped at random.
* 10% Condition: This sample of 348 truckers is less than 10% of all truckers.
* Success/Failure: 49 tickets and 299 tickets are both at least 10, so our sample is large enough.

Under these conditions the sampling distribution of the proportion can be modeled by a Normal model.

We will find a one-proportion $z$-interval.

We know $n = 348$ and $\hat{p} = 0.14$, so 
$$SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(0.14)(.86)}{348}} = 0.0186$$

The sampling model is Normal, for a 95% confidence interval the critical value is $z^* = 1.96$.

The margin of error is $ME = z^* \times SE(\hat{p}) = 1.96(0.0186) = 0.0365$.

The 95% confidence interval is $0.14 \pm 0.0365$ or $(0.1035, 0.1765)$.

We are 95% confident that between 10.4% and 17.7% of truck drivers have driven too many hours in a day.

If we repeated the sampling and created new confidence intervals many times we would expect about 95% of those intervals to contain the actual proportion of truck drivers that have driven too many hours in a day.