

Midterm Exam Review Questions

Due No due date **Points** 30 **Questions** 30 **Time Limit** None

Instructions

This is a set of practice questions that covers material from Lectures 1-14. You are welcome to work with others to answer the questions.

Completion of these questions is worth 5% extra credit on your overall midterm exam score (up to a maximum score of 100%), so you can earn extra credit in advance before taking the exam! For example, if you get 80% of the practice questions correct, you will still earn the 5% extra credit. After you complete all the practice questions, you can see the correct answers and review your responses to the questions.

If unstated, assume the significance level (alpha) is 5%.

If you notice any typos or have questions about an answer please let Alex know via Canvas message, email, or a post on the discussion board.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	162 minutes	22 out of 30

Submitted Oct 14 at 6:37pm

Correct!	Question 1	1 / 1 pts
	The sample mean is an unbiased estimator of the population mean.	
	<input checked="" type="radio"/> True	
	<input type="radio"/> False	
	<div style="border: 1px solid #ccc; padding: 5px;">True. See Lecture 2 (CLT), Page 7, for the derivation.</div>	

Question 2	1 / 1 pts
-------------------	------------------

Iowans love their corn, it's just a fact! Some intrepid researchers did some research into the number of ears of corn on the cob that a typical Iowan eats for breakfast during peak sweet corn season and summarized the results in the following table:

# of Ears	0	1	2	3	4
P(X=x)	0.1	0.5	0.2	0.15	0.05

What is the expected value of number of ears of corn eaten for breakfast? Round your answer to two decimal places.

Correct!

1.55

Correct Answers

1.55 (with margin: 0)

$$E(X) = 0(0.1) + 1(0.5) + 2(0.2) + 3(0.15) + 4(0.05) = 0 + 0.5 + 0.4 + 0.45 + 0.2 = 1.55$$

Question 3

0 / 1 pts

Iowans love their corn, it's just a fact! Some intrepid researchers did some research into the number of ears of corn on the cob that a typical Iowan eats for breakfast during peak sweet corn season and summarized the results in the following table:

# of Ears	0	1	2	3	4
P(X=x)	0.1	0.5	0.2	0.15	0.05

What is the standard deviation of ears of corn eaten for breakfast? Round your answer to two decimal places.

You Answered

0.77

Correct Answers

1.02 (with margin: 0)

One method is to take the variance formula directly:

$$V(X) = \sum_x (x - \mu)^2 p(x) = \sigma^2$$

$$\begin{aligned} V(X) &= (0 - 1.55)^2(0.1) + (1 - 1.55)^2(0.5) + (2 - 1.55)^2(0.2) + (3 - 1.55)^2(0.15) \\ &\quad + (4 - 1.55)^2(0.05) = 1.0475 \end{aligned}$$

Then we take the square root for our SD: $\sqrt{1.0475} = 1.023474$

Alternatively you could also calculate $E(X^2)$ and use

$$V(X) = E(X^2) - E(X)^2$$

Question 4

1 / 1 pts

A study recently came out that suggests an overwhelming 90% of people prefer raisins in their cookies (source: The Raisin-America Industry Society for Innovative Nourishment journal). You decide to take a random sample of 10 individuals. What is the probability that two or fewer people prefer raisins?

- P(X=0) + P(X=1)
- P(X=0) + P(X=1) + P(X=2)
- P(X=3) + P(X=4) + ... + P(X=10)
- P(X=2) + P(X=3) + ... + P(X=10)

Correct!

P(X=0) + P(X=1) + P(X=2) is the correct answer. We could use the binomial distribution to then determine the probability of having two or fewer people in our random sample prefer raisins, based on the new research conducted by the RAISIN journal, is

$${10 \choose 0}(0.9)^0(0.1)^1 + {10 \choose 1}(0.9)^1(0.1)^9 + {10 \choose 2}(0.9)^2(0.1)^8 = 0.0000003736$$

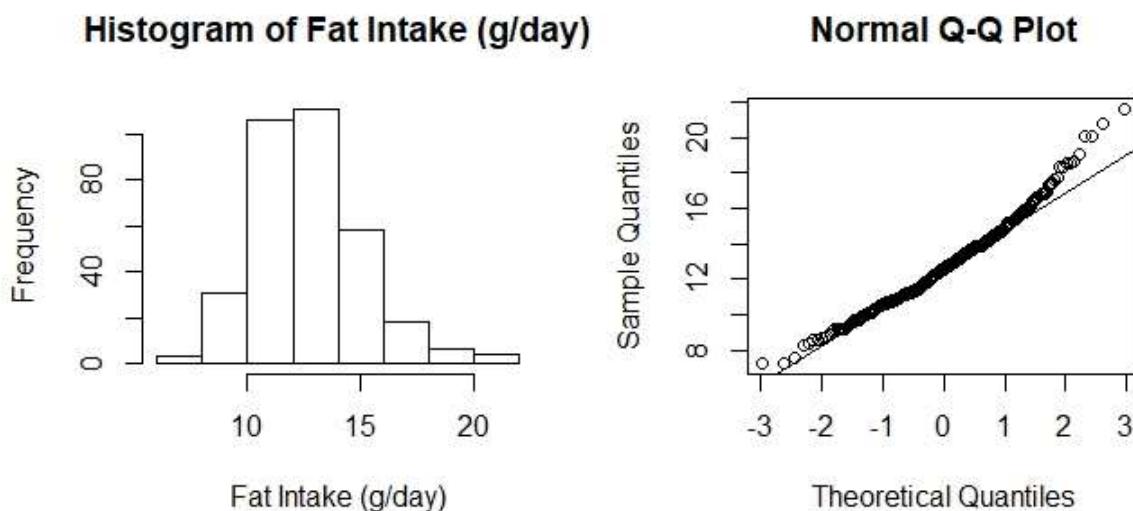
In other words, very, very unlikely if the 90% of raisin preference is true.

Question 5

0 / 1 pts

The Epi package contains a set of example data called *diet* looking at a subsample of 337 subjects from a larger cohort study on the incidence of coronary heart disease. Assume a researcher is interested in using this data to test a hypothesis about fat intake (measured in grams/day) being different for people who are under vs. over 5'10" (approximately 177.8 cm). The first question they have is if their fat intake is normally distributed. Based on the figures and Shapiro-Wilk test below, what would you say:

Shapiro-Wilk test has $W=0.97734$, $p=0.00003$



Correct Answer

The evidence is mixed and we should carefully consider if the normal distribution can be assumed

Definitely normally distributed

Definitely NOT normally distributed

You Answered

Why waste time with figures? $N=337 \gg 30$, CLT will save us all!

Histogram is approximately normal with perhaps some right skew.

QQ plot suggests potential deviation from normality in the tails.

The Shapiro-Wilk test suggests normality is violated.

Overall this might suggest caution should be taken. *However*, on an exam where you can write out your thought process you can really justify the response on a continuum:

- Shapiro-Wilk is significant, but we know it can produce small p-values for larger samples even if the data is nearly normal.
- The skewness to the histogram is somewhat subtle, perhaps more bars would help identify the distribution even better.
- The QQ plot does deviate in the tails, but this matches the slight right skewness of the histogram and the bulk of data is normal.

The one answer that isn't correct is throwing all caution to the wind and using the CLT, since we also need to consider the accuracy we wish to achieve.

Question 6

1 / 1 pts

The Epi package contains a set of example data called *diet* looking at a subsample of 337 subjects from a larger cohort study on the incidence of coronary heart disease. Assume a researcher is interested in using this data to test a hypothesis about fat intake (measured in grams/day) being different for people who are under vs. over 5'10" (approximately 177.8 cm). What is the appropriate null hypothesis they should specify (if they are assuming normality)?

-
- The mean for those under 5'10" is less than the mean for those over 5'10"
 - The modes are equal for our two groups.
 - The means are not equal for our two groups
 - The means are equal for our two groups.
-

Correct!

The correct null hypothesis, given that we are comfortable assuming normality, is the the means between the two groups are equal. Alternatively, we could also state that the difference between the mean fat intake (g/day) of our two groups is 0.

Question 7

1 / 1 pts

The Epi package contains a set of example data called *diet* looking at a subsample of 337 subjects from a larger cohort study on the incidence of coronary heart disease. Assume a researcher is interested in using this data to test a hypothesis about fat intake (measured in grams/day) being different for people who are under vs. over 5'10" (approximately 177.8 cm). They conduct a two-sample t-test (where those under 5'10" are group TRUE) and produce the following results. What is your interpretation if we assume $\alpha = 0.05$?

```
Welch Two Sample t-test

data: diet$fat by diet$f10
t = 2.4239, df = 156.2, p-value = 0.0165
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1330229 1.3044568
sample estimates:
mean in group FALSE   mean in group TRUE
 13.29321           12.57447
```

Correct!

- p<0.05, reject H₀ that the mean fat intake is the same in our two groups.
- p<0.05, fail to reject H₀ that the mean fat intake is the same in our two groups.
- p<0.05, fail to accept H₀ that the mean fat intake is the same in our two groups.
- p<0.05, accept H₀ that the mean fat intake is the same in our two groups.

Since we assume alpha=0.05, p=0.0165 is less than this threshold and we reject H₀. Further, we could note that based on the estimates provided by the t.test function in R, the group under 5'10" has a lower mean fat intake in g/day of 12.6 compared to 13.3 for those over 5'10".

Also recall, we never "accept" the null hypothesis. We only ever have evidence to reject or fail to reject it.

Question 8

1 / 1 pts

The Epi package contains a set of example data called *diet* looking at a subsample of 337 subjects from a larger cohort study on the incidence of coronary heart disease. Assume a researcher is interested in using this data to test a hypothesis about fat intake (measured in grams/day) being different for people who are under vs. over 5'10" (approximately 177.8 cm). If you ultimately were unwilling to assume normality and use a two-sample t-test, what would be an alternative test you could use (choose all appropriate options listed)?

Two-Sample Z-test

Permutation Test

Chi-Squared Test

Correct!

Wilcoxon Rank Sum

Fisher's Exact Test

One-Sample t-Test

Correct!

Of the options provided, two potential choices would be the non-parametric Wilcoxon rank sum (also known as the Mann Whitney U) test or to implement a permutation test. The other choices are either for categorical data or still assume normality.

Question 9

1 / 1 pts

Dr. Pine recently received a small pilot award to conduct a study in 16 participants for a drug they are developing to reduce LDL cholesterol. A clinically meaningful difference would be a change of 8 mg/dL in LDL-C.

The standard deviation for change is 10 mg/dL. Assuming $\alpha=0.05$ and 80% power for a two-sided test, what is their detectable difference? (Use the reference table below, round answer to two decimal places.)

For reference, use:

Z _{0.80}	Z _{0.85}	Z _{0.90}	Z _{0.95}	Z _{0.975}	Z _{0.99}	Z _{0.999}
0.84	1.04	1.28	1.64	1.96	2.33	3.09

Correct!

7

Correct Answers

7 (with margin: 0)

Detectable Difference =

$$(Z_{1-\beta} + Z_{1-\alpha/2})\sigma/\sqrt{n} = (0.84 + 1.96)(10/4) = 7$$

This suggests that the study is able to detect at least the clinically meaningful difference of 8 given the assumptions provided by Dr. Pine.

Question 10

1 / 1 pts

Dr. Cedar wants to apply for a grant from the NIH for a basic science study in gerbils for hearing loss (measured in decibels of hearing loss, dB HL) who are exposed to chronic background noise. They are unsure of how many gerbils they should request, but they note that in a pilot study the standard deviation for change was 100 dB HL and that a change of 40 dB HL would define their threshold for clinical intervention. They want to have $\alpha=0.10$ and 90% power. What sample size of gerbils do they need for a two-sided test?

For reference, use:

Z _{0.80}	Z _{0.85}	Z _{0.90}	Z _{0.95}	Z _{0.975}	Z _{0.99}	Z _{0.999}
0.84	1.04	1.28	1.64	1.96	2.33	3.09

Correct!

54

Correct Answers

54 (with margin: 0)

$$n = \frac{\sigma^2(Z_{1-\beta}+Z_{1-\alpha/2})^2}{(\mu_0-\mu_1)^2} = \frac{100^2(1.28+1.64)^2}{40^2} = 53.29$$

Recall, for sample size we must always round *up* to the nearest whole number to ensure we maintain adequate power for our given assumptions. So ultimately they need 54 gerbils.

Question 11

1 / 1 pts

Dr. Cedar is a bit surprised by the number of gerbils in your last answer. You mention that it sounds like their hypothesis might actually need a one-sided sample size calculation since hearing loss can't improve in their study (i.e. no intervention to improve hearing). What would the needed sample size be for a one-sided test with the same assumptions of the standard deviation for change was 100 dB HL and that a change of 40 dB HL would define their threshold for clinical intervention with $\alpha=0.10$ and 90% power?

For reference, use:

Z _{0.80}	Z _{0.85}	Z _{0.90}	Z _{0.95}	Z _{0.975}	Z _{0.99}	Z _{0.999}
0.84	1.04	1.28	1.64	1.96	2.33	3.09

Correct!

41

Correct Answers

41 (with margin: 0)

$$n = \frac{\sigma^2(Z_{1-\beta}+Z_{1-\alpha})^2}{(\mu_0-\mu_1)^2} = \frac{100^2(1.28+1.28)^2}{40^2} = 40.96$$

Recall, for sample size we must always round *up* to the nearest whole number to ensure we maintain adequate power for our given assumptions.
So ultimately they need 41 gerbils.

Question 12

1 / 1 pts

Celiac disease is an autoimmune disease where eating gluten (a protein found in wheat, barley, and rye) leads to an immune response in your small intestine. The symptoms of celiac vary from "silent" (i.e., no observable or noticeable symptoms even though damage is still being done to the small intestine) to severe (i.e., miscarriage, hospitalization, anemia, diarrhea). The only treatment is a strict gluten-free diet once celiac has been diagnosed.

The gold standard is an endoscopy of the small intestine for biopsy samples of the villi, but this can be invasive and uncomfortable. Blood tests exist, but it can take days for results. The Redwood-Young Environmental Labs (RYE Labs) developed a new test with nearly immediately results based on a finger prick. After going through IRB approval, they team up with a GI clinic in Kansas to compare their test to the gold standard. Their method correctly identified 5 out of 15 patients with a positive endoscopy and falsely identified 15 out of 85 patients.

What is the sensitivity of the test? (Round to the nearest whole percent.)

Correct!

33

Correct Answers

33 (with margin: 1)

0.33 (with margin: 0.01)

The 2x2 table we can create from their study is:

	Endoscopy +	Endoscopy -
RYE +	5	15
RYE -	10	70

The sensitivity is $a/(a+c) = 5/(5+10) = 5/15 = 0.33$, or 33%.

Question 13

1 / 1 pts

Celiac disease is an autoimmune disease where eating gluten (a protein found in wheat, barley, and rye) leads to an immune response in your small intestine. The symptoms of celiac vary from "silent" (i.e., no observable or noticeable symptoms even though damage is still being done to the small intestine) to severe (i.e., miscarriage, hospitalization, anemia, diarrhea). The only treatment is a strict gluten-free diet once celiac has been diagnosis.

The gold standard is an endoscopy of the small intestine for biopsy samples of the villi, but this can be invasive and uncomfortable. Blood tests exist, but it can take days for results. The Redwood-Young Environmental Labs (RYE Labs) developed a new test with nearly immediate results based on a finger prick. After going through IRB approval, they team up with a GI clinic in Kansas to compare their test to the gold standard. Their method correctly identified 5 out of 15 patients with a positive endoscopy and falsely identified 15 out of 85 patients.

What is the specificity of the test? (Round to the nearest whole percent.)

Correct!

82

Correct Answers

82 (with margin: 1)

0.82 (with margin: 0.01)

Specificity = $d/(b+d) = 70/(15+70) = 70/85 = 0.82$ or 82%

Question 14

1 / 1 pts

Celiac disease is an autoimmune disease where eating gluten (a protein found in wheat, barley, and rye) leads to an immune response in your small intestine. The symptoms of celiac vary from "silent" (i.e., no observable or noticeable symptoms even though damage is still being done to the small intestine) to severe (i.e., miscarriage, hospitalization, anemia, diarrhea). The only treatment is a strict gluten-free diet once celiac has been diagnosis.

The gold standard is an endoscopy of the small intestine for biopsy samples of the villi, but this can be invasive and uncomfortable. Blood tests exist, but it can take days for results. The Redwood-Young Environmental Labs (RYE Labs) developed a new test with nearly immediately results based on a finger prick. After going through IRB approval, they team up with a GI clinic in Kansas to compare their test to the gold standard. Their method correctly identified 5 out of 15 patients with a positive endoscopy and falsely identified 15 out of 85 patients.

What is the PPV? (Round to the nearest whole percent.)

Correct!

25

Correct Answers

0.25 (with margin: 0.01)
25 (with margin: 1)

$$\text{PPV} = a/(a+b) = 5/(5+15) = 5/20 = 0.25, \text{ or } 25\%.$$

Note, we assume the prevalence from our sample in the 2x2 table is accurate.

Question 15

1 / 1 pts

Celiac disease is an autoimmune disease where eating gluten (a protein found in wheat, barley, and rye) leads to an immune response in your small intestine. The symptoms of celiac vary from "silent" (i.e., no observable or noticeable symptoms even though damage is still being done to the small intestine) to severe (i.e., miscarriage, hospitalization, anemia, diarrhea). The only treatment is a strict gluten-free diet once celiac has been diagnosis.

The gold standard is an endoscopy of the small intestine for biopsy samples of the villi, but this can be invasive and uncomfortable. Blood tests exist, but it can take days for results. The Redwood-Young Environmental Labs (RYE Labs) developed a new test with nearly immediately results based on a finger prick. After going

through IRB approval, they team up with a GI clinic in Kansas to compare their test to the gold standard. Their method correctly identified 5 out of 15 patients with a positive endoscopy and falsely identified 15 out of 85 patients.

What is the NPV? (Round to the nearest whole percent.)

Correct!

88

Correct Answers

88 (with margin: 1)
0.88 (with margin: 0.01)

$$NPV = d/(c+d) = 70 / (10+70) = 70/80 = 0.88 \text{ or } 88\%$$

Note, assuming the prevalence of the sample is the true prevalence.

Question 16

1 / 1 pts

Celiac disease is an autoimmune disease where eating gluten (a protein found in wheat, barley, and rye) leads to an immune response in your small intestine. The symptoms of celiac vary from "silent" (i.e., no observable or noticeable symptoms even though damage is still being done to the small intestine) to severe (i.e., miscarriage, hospitalization, anemia, diarrhea). The only treatment is a strict gluten-free diet once celiac has been diagnosis.

The gold standard is an endoscopy of the small intestine for biopsy samples of the villi, but this can be invasive and uncomfortable. Blood tests exist, but it can take days for results. The Redwood-Young Environmental Labs (RYE Labs) developed a new test with nearly immediately results based on a finger prick. After going through IRB approval, they team up with a GI clinic in Kansas to compare their test to the gold standard. Their method correctly identified 5 out of 15 patients with a positive endoscopy and falsely identified 15 out of 85 patients.

What is the LR+ of the test? (Round to 2 decimal places.)

Correct!

1.89

Correct Answers

1.83 (with margin: 0.1)

$$LR+ = Se/(1-Sp) = 0.33/(1-0.82) = 1.83$$

Question 17

1 / 1 pts

Celiac disease is an autoimmune disease where eating gluten (a protein found in wheat, barley, and rye) leads to an immune response in your small intestine. The symptoms of celiac vary from "silent" (i.e., no observable or noticeable symptoms even though damage is still being done to the small intestine) to severe (i.e., miscarriage, hospitalization, anemia, diarrhea). The only treatment is a strict gluten-free diet once celiac has been diagnosis.

The gold standard is an endoscopy of the small intestine for biopsy samples of the villi, but this can be invasive and uncomfortable. Blood tests exist, but it can take days for results. The Redwood-Young Environmental Labs (RYE Labs) developed a new test with nearly immediately results based on a finger prick. After going through IRB approval, they team up with a GI clinic in Kansas to compare their test to the gold standard. Their method correctly identified 5 out of 15 patients with a positive endoscopy and falsely identified 15 out of 85 patients.

What is the LR- of the test? (Round to 2 decimal places.)

Correct!

0.81

Correct Answers

0.82 (with margin: 0.1)

$$LR- = (1-Se)/Sp = (1-0.33)/0.82 = 0.82$$

Question 18

1 / 1 pts

Based on your summaries of RYE's study for their finger prick test to rapidly detect celiac disease, is it a good test?

Correct!

Yes, the sensitivity suggests 33% of people with celiac will test positive. It's specificity suggests that only 82% of those without celiac will test negative (i.e., 18% false positives).

Correct!

No, the sensitivity suggests only 33% of people with celiac will test positive. It's specificity suggests that only 82% of those without celiac will test negative (i.e., 18% false positives).

The sensitivity of this test is not great at 33%, and the specificity of 82% suggests lots of false positives. Additionally we can discuss the PPV being only 25%, missing many of those with the disease, or the likelihood ratios being near 1.

Question 19

1 / 1 pts

Assume that RYE Labs was interested in estimating the posterior probability of the test in populations that aren't entirely surrounded by wheat fields and perhaps are less likely to have celiac disease. As a hail mary, they take the national US estimate that 1% of individuals have celiac disease. What would the posterior probability of having celiac be, given you had a positive test, when considering the entire US? (Round to the nearest whole percent.)

Correct!

2

Correct Answers

2 (with margin: 1)

0.02 (with margin: 0.01)

Recall the relationship between PPV and the positive posterior probability...they're the same! So we can cut straight to the chase by using our more detailed PPV equation that doesn't rely on the sample cells:

Posterior prob of celiac given a positive test = PPV =

$$\frac{Se \times P(D)}{Se \times P(D) + (1 - Sp) \times (1 - P(D))} = \frac{0.33(0.01)}{0.33(0.01) + (1 - 0.82)(1 - 0.01)} = .0182$$

Rounding to the nearest percent, we have a posterior probability of having celiac given the national prevalence of 1% that is only 2%!

Note, we could also have gone the long way:

$$\text{Prior odds of } D = P(D)/(1-P(D)) = 0.01 / (1-0.01) = 0.01 / 0.99 = 0.0101$$

$$\text{Posterior odds of } D = (.01/.99) \times LR+ = (.01/.99) \times (Se/(1-Sp)) = (.01/.99) \times (.33/(1-.82)) = 0.0185$$

$$\text{Posterior prob of } D = 0.0185 / (1 + 0.0185) = 0.0182$$

Question 20

1 / 1 pts

Aronsson, et al., examined a prospective cohort study that was examining potential environmental determinants in developing diabetes in youths (Aronsson, Carin Andrén, et al. "Age at gluten introduction and risk of celiac disease." *Pediatrics* (2015)), to determine if age of introduction to gluten was protective or a risk factor.

For our purposes, we will work with a subset of their data:

	Develop Celiac	Don't Develop Celiac
Exposed to Gluten <17 weeks	13	383
Exposed to Gluten 17-26 weeks	141	2152

Letting p_1 be the 17-26 week group and p_2 be the <17 week group, what is the risk difference between age groups (round to the nearest whole percent).

Correct!

3

Correct Answers

0.03 (with margin: 0.01)

3 (with margin: 1)

The RD = $p_1 - p_2 = (141/2293) - (13/396) = 0.03$ or 3%.

Question 21

0 / 1 pts

Aronsson, et al., examined a prospective cohort study that was examining potential environmental determinants in developing diabetes in youths (Aronsson, Carin Andrén, et al. "Age at gluten introduction and risk of celiac disease." *Pediatrics* (2015)), to determine if age of introduction to gluten was protective or a risk factor.

For our purposes, we will work with a subset of their data:

	Develop Celiac	Don't Develop Celiac
Exposed to Gluten <17 weeks	13	383
Exposed to Gluten 17-26 weeks	141	2152

Letting p_1 be the 17-26 week group and p_2 be the <17 week group, what is the risk ratio between age groups (round to two decimals).

You Answered

0.53

Correct Answers

1.87 (with margin: 0.05)

$p_1/p_2 = (141/2293) / (13/396) = 1.873126$

Question 22

0 / 1 pts

Aronsson, et al., examined a prospective cohort study that was examining potential environmental determinants in developing diabetes in youths (Aronsson, Carin Andrén, et al. "Age at gluten introduction and risk of celiac

disease." *Pediatrics* (2015)), to determine if age of introduction to gluten was protective or a risk factor.

For our purposes, we will work with a subset of their data:

	Develop Celiac	Don't Develop Celiac
Exposed to Gluten <17 weeks	13	383
Exposed to Gluten 17-26 weeks	141	2152

Letting p_1 be the 17-26 week group and p_2 be the <17 week group, what is the odds ratio between age groups (round to two decimals).

You Answered

0.52

Correct Answers

1.93 (with margin: 0.05)

$$OR = (p_1/(1-p_1)) / (p_2/(1-p_2)) = 1.930333$$

Question 23

0 / 1 pts

Aronsson, et al., examined a prospective cohort study that was examining potential environmental determinants in developing diabetes in youths (Aronsson, Carin Andrén, et al. "Age at gluten introduction and risk of celiac disease." *Pediatrics* (2015)), to determine if age of introduction to gluten was protective or a risk factor.

For our purposes, we will work with a subset of their data:

	Develop Celiac	Don't Develop Celiac
Exposed to Gluten <17 weeks	13	383
Exposed to Gluten 17-26 weeks	141	2152

Letting p_1 be the 17-26 week group and p_2 be the <17 week group, calculate the 95% confidence interval around the RR. What can we conclude?

You Answered



The CI contains 1, so we cannot determine that age of gluten exposure is associated to developing celiac disease.

Correct Answer

The CI doesn't contain 1, so waiting for exposure until 17-26 weeks relative to <17 weeks is a risk factor to developing celiac.

The 95% CI for RR is determined as:

$$\exp(\log(RR) \pm 1.96 \times SE(\log(RR))) = \exp(\log(1.87) \pm 1.96 \times \sqrt{\frac{2152}{141 \times 2293} + \frac{383}{13 \times 396}}) = (1.072, 3.272)$$

Question 24

0 / 1 pts

From a pilot study sample of 20 subjects, an 80% two-sided confidence interval for mean weight loss (% body weight lost from baseline) during an experimental twelve-week diet + Tai Chi program is found to be (1.5%, 7.5%). % Body Weight Lost = $\{(\text{Baseline Weight} - \text{Ending Weight}) / \text{Beginning Weight}\} \times 100\%$, so it's a continuous measure of weight loss that could theoretically range from a large negative % to a large positive %.

Using the results provided without doing any additional work, if a 5% loss in twelve weeks is considered to be of **clinical** benefit, do we have clinical significance and are the pilot results encouraging in terms of further study with more subjects?

[Note: Fall 2019 BIOS 6611, we did not discuss clinical vs. statistical significance in detail this semester. So this question of *clinical* significance will not be on your midterm exam, but is still useful to consider and be exposed to.]

Definitely NO clinical benefit, the confidence interval includes values below the 5% benefit and it would be a waste of resources to do a larger study.

You Answered

Definite clinical benefit since 5% is included in our confidence interval. A follow-up study is not needed since these results are conclusive.

Correct Answer

The results are inconclusive. The confidence interval includes the clinically beneficial value, but also includes values that are not of clinical significance. A follow-up study with more subjects would enable us to more precisely determine clinical benefit.

When considering the **clinical** significance of these results we can note that the 80% confidence interval includes our value of clinical significance *and* values that are less than clinically significant. In this case a follow-up study may be warranted with more participants so we can get a narrower estimate of the interval.

Note: In this case we are presented with an 80% confidence interval instead of our more typical 95% confidence interval. We know that increasing our confidence interval from 80% to 95% will result in a wider interval. However these wider intervals will still contain our clinically significant value of 5% indicating an inconclusive clinical significance.

Question 25

0 / 1 pts

From a pilot study sample of 20 subjects, an 80% two-sided confidence interval for mean weight loss (% body weight lost from baseline) during an experimental twelve-week diet + Tai Chi program is found to be (1.5%, 7.5%). % Body Weight Lost = $\{(\text{Baseline Weight} - \text{Ending Weight}) / \text{Beginning Weight}\} \times 100\%$, so it's a continuous measure of weight loss that could theoretically range from a large negative % to a large positive %.

Using the results provided without doing any additional work, do the results suggest **statistical significance** with respect to a null hypothesis of no change?

- We cannot determine statistical significance from the information provided.

- Yes, assuming $\alpha=0.20$ our results are statistically significant.

- Yes, assuming $\alpha=0.05$ our results are statistically significant.

- No, our results are not statistically significant at $\alpha=0.20$.

Correct Answer

You Answered

In this case we are given the 80% confidence interval of (1.5%, 7.5%). Since the confidence interval does not include 0% it is statistically significant. We have to be careful, however, because we aren't provided our traditional 95% CI with $\alpha=0.05$. In this case the 80% CI corresponds to $\alpha =0.20$.

Question 26

0 / 1 pts

Based on the previous Tai Chi pilot study we can derive that the mean weight loss was 4.5% (the mean of the lower and upper confidence interval values). Further, we can work backwards from either bound of our CI to solve for the standard deviation, given that we know the sample size was 20:

$$4.5 + z_{1-0.2/2} \frac{\sigma}{\sqrt{20}} = 4.5 + 1.28 \frac{\sigma}{\sqrt{20}} = 7.5$$

Solving for sigma we find that $\sigma = \frac{(7.5-4.5)}{1.28} \times \sqrt{20} = 10.48$

The researchers decide to proceed with a slightly larger study of 30 individuals. They calculate their power for a two-sided test to achieve their clinically meaningful weight loss target given alpha=0.05 and n=30 to be

$$1 - \beta = \Phi \left[\frac{5}{10.48/\sqrt{30}} - Z_{1-0.05/2} \right] = \Phi(0.653) = 0.743$$

What is the meaning of power in the context of their study?



Given the null hypothesis of 0% weight loss is true, the probability of incorrectly rejecting H0 of 0% weight loss is 74.3%



The probability of making a type I error is 1-0.743, or 25.7%.



We are 74.3% likely to detect the alternative hypothesis of 5% weight loss.

You Answered

Correct Answer



Given the null hypothesis of 0% weight loss is false, the probability of rejecting the null is 74.3% given our desired detectable difference of 5%.

Question 27

1 / 1 pts

Dr. Hera Ade is conducting a study on the potential ototoxicity due to the accumulated exposure to a commonly used antibiotic. They are concerned that some people may have a genetic susceptibility to increased toxicity, and that if this could be identified we could modify treatment to use less toxic interventions to avoid hearing loss. One SNP in particular has been identified as a potential candidate for ototoxicity of this antibiotic. Given the early stage of their research hypothesis, they conducted an animal study in ferrets with homozygous dominant, heterozygous, and homozygous recessive variants of the SNP. After equal exposure to the antibiotic, a startle test is conducted with respect to a sound trigger and the results (in seconds) are as follows:

	Homozygous Dominant (n=45)	Heterozygous (n=42)	Homozygous Recessive (n=44)
Mean	5.3	4.3	6.9
SD	3.0	3.5	5.8

Based on these results, assuming all assumptions are met and a significance level of 0.05, interpret the outcomes of a one-way ANOVA assuming equal variances with a p-value of 0.0246.

Since $p < 0.05$, we reject the null hypothesis that the group means are all equal. The homozygous recessive is significantly longer than the other two.

With three groups, our significance level should be something like $0.05/3=0.0167$. Since $p=0.0246$ is not less than 0.0167, we fail to reject the null hypothesis that all means are equal.

Since $p < 0.05$, we fail to reject the null hypothesis that the group means are all equal.

Correct!

Since $p < 0.05$, we reject the null hypothesis that the group means are all equal. At least one of the means is significantly different.

If we assume that all assumptions are met, then our one-way ANOVA tests the null hypothesis that all group means are equal. However, without post-hoc testing, we cannot state *which* groups are different, only that at least one is different.

For our observed p-value of 0.0246 and significance level of 0.05, we would reject H₀ and state the at least one group is different.

Note, we do not need to correct for multiple groups since the the null hypothesis is one test for at least 1 group being different. With post-hoc tests, we would want to adjust for multiple comparisons to protect against type I errors.

Question 28

1 / 1 pts

Dr. Hera Ade is conducting a study on the potential ototoxicity due to the accumulated exposure to a commonly used antibiotic. They are concerned that some people may have a genetic susceptibility to increased toxicity, and that if this could be identified we could modify treatment to use less toxic interventions to avoid hearing loss. One SNP in particular has been identified as a potential candidate for ototoxicity of this antibiotic. Given the early stage of their research hypothesis, they conducted an animal study in ferrets with homozygous dominant, heterozygous, and homozygous recessive variants of the SNP. After equal exposure to the antibiotic, a startle test is conducted with respect to a sound trigger and the results (in seconds) are as follows:

	Homozygous Dominant (n=45)	Heterozygous (n=42)	Homozygous Recessive (n=44)
Mean	5.3	4.3	6.9
SD	3.0	3.5	5.8

While assuming all assumptions are met can make analyses "easier," you have concerns based on the summary statistics in the table above that some of these assumptions may not be met. Of immediate concern, the assumption of homogeneous variances seems suspect given the observed data. You conduct Bartlett's test and observe p=0.00002. What do you conclude?

Correct!

Since $p < 0.05$, we fail to reject the null hypothesis of equal variances for all groups.

Since $p > 0.05$, we fail to reject the null hypothesis of equal variances for all groups.

Since $p < 0.05$, we reject the null hypothesis of equal variances for all groups. At least one group has a different variance.

Since $p > 0.05$, we reject the null hypothesis of equal variances for all groups. At least one group has a different variance.

Bartlett's test for homogeneity of variance can be used to test the assumption of equal variances across groups for the one-way ANOVA. The null hypothesis is that the variances are equal across all groups, whereas the alternative is that at least one group is different (and therefore the standard one-way ANOVA assuming equal variances is inappropriate).

For our problem, $p < 0.001$, so we reject H_0 that the variances are equal across the 3 groups. Therefore, we should use a more appropriate method to compare the groups.

Question 29

1 / 1 pts

Dr. Hera Ade is conducting a study on the potential ototoxicity due to the accumulated exposure to a commonly used antibiotic. They are concerned that some people may have a genetic susceptibility to increased toxicity, and that if this could be identified we could modify treatment to use less toxic interventions to avoid hearing loss. One SNP in particular has been identified as a potential candidate for ototoxicity of this antibiotic. Given the early stage of their research hypothesis, they conducted an animal study in ferrets with homozygous dominant, heterozygous, and homozygous recessive variants of the SNP. After equal exposure to the antibiotic, a startle test is conducted with respect to a sound trigger and the results (in seconds) are as follows:

	Homozygous	Heterozygous	Homozygous
--	------------	--------------	------------

	Dominant (n=45)	(n=42)	Recessive (n=44)
Mean	5.3	4.3	6.9
SD	3.0	3.5	5.8

Based on the results of your Bartlett's test, you decide that Welch's ANOVA, which doesn't assume equal variances, should be used instead. For Welch's ANOVA you observe $p=0.053$. What is your conclusion (assuming a significance level of 0.05)?

Correct!

Since $p>0.05$, we fail to reject the null hypothesis that all group means are equal.

Since $p<0.05$, we fail to reject the null hypothesis that all group means are equal.



Since $p>0.05$, we reject the null hypothesis that at least one group mean is different.



Since $p<0.05$, we reject the null hypothesis that all group means are equal and conclude that at least one is different..

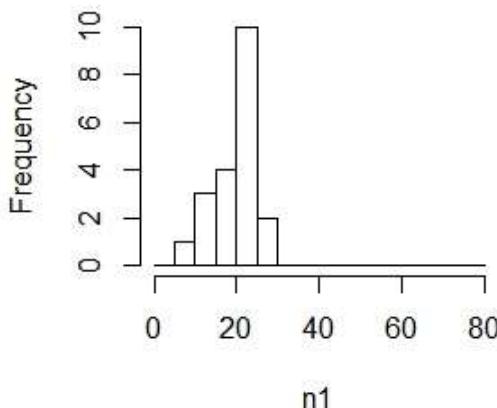
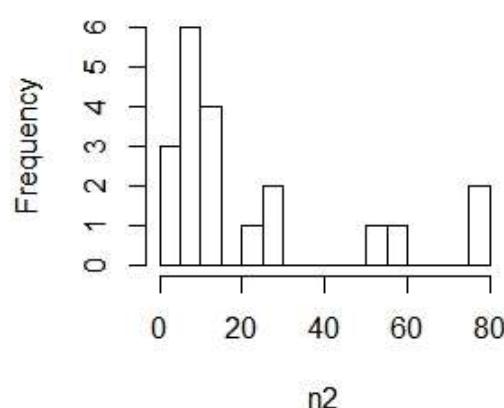
By using the more appropriate Welch's ANOVA that accommodates unequal variances, our $p=0.053$ is above our significance level of 0.05. With this result, we fail to reject the null hypothesis that all group means are equal.

Therefore, the startle response in seconds for our different genetic groups of ferrets are not statistically different at the 0.05 significance level.

Question 30

1 / 1 pts

A colleague conducts an experiment with 20 subjects in two different groups. They plot a histogram of the results and are concerned about assuming normality, so they conduct a Wilcoxon rank sum (aka Mann-Whitney U test) and plot a histogram of their data. The Wilcoxon rank sum test has a p-value of 0.127. What is the appropriate conclusion from the provided options?

Group 1 Histogram**Group 2 Histogram**

Correct!



Since $p>0.05$, we fail to reject the null hypothesis that the mean ranks are equal between the two groups.



Since $p>0.05$, we fail to reject the null hypothesis that the medians are equal between groups.



Since $p>0.05$, we fail to reject the null hypothesis that the means are the same between groups.

Under special circumstances, we are able to assume that the Wilcoxon rank sum/Mann-Whitney/Wilcoxon-Mann-Whitney test is actually a test of the medians (i.e., we assume the distributions are the same under the null and we are actually examining a shift in the location). In general, we are safe to always state that we are testing the null hypothesis of mean ranks being equal between groups because that is technically the null hypothesis of these tests (however, this is usually unsatisfying since it doesn't really map directly to the raw data someone has collected).