STATS 8 w19 stats8-w19-finalvB

Brian Nhan Thien Chung

TOTAL POINTS

75.5 / 92

QUESTION 1

Physical activity - resistance training (Corey Katz) 12 pts

1.1 confidence interval (Corey Katz) 3 / 3

✓ **- 0 pts Correct**

1.2 margin of error (Corey Katz) 2 / 2

✓ **- 0 pts Correct**

1.3 interpretation (Corey Katz) 2.5 / 4

✓ **- 0.5 pts Missing units**

✓ **- 1 pts Missing Population that we are conducting inference on, or did not specify the correct population**

1.4 hypothesis from CI (Corey Katz) 3 / 3

✓ **- 0 pts Correct**

QUESTION 2

Physical activity - resistance vs endurance (Corey Katz) 10 pts

2.1 hypotheses (Corey Katz) 3 / 3

✓ **- 0 pts Correct**

2.2 P-value (Corey Katz) 3 / 3

✓ **- 0 pts Correct**

2.3 interpretation (Corey Katz) 1.5 / 4

✓ **- 1 pts no population defined**

✓ **- 1.5 pts not complete**

QUESTION 3

Fluoroquinolone (Chu Ching) 9 pts

3.1 birth defect (Chu Ching) 3 / 3

✓ **+ 3 pts correct**

3.2 head circumference (Chu Ching) 3 / 3

✓ **+ 3 pts correct**

3.3 maternal age (Chu Ching) 3 / 3

✓ **+ 3 pts correct**

QUESTION 4

YouTube influencers (Shannon Proctor) 11 pts

4.1 hypotheses (Shannon Proctor) 3 / 3

✓ **- 0 pts Correct**

4.2 interpretation (Shannon Proctor) 3 / 4

✓ **- 1 pts Made conclusion based on alternative hypothesis.**

We NEVER conclude that we "favor" the alternative hypothesis, we can only reject or fail to reject the null hypothesis.

4.3 which condition? (Chu Ching) 2 / 2

✓ **+ 2 pts correct**

4.4 conditions met? (Chu Ching) 2 / 2

✓ **+ 2 pts correct**

QUESTION 5

Chicken phenotypes (Corey Katz) 7 pts

5.1 hypotheses (Corey Katz) 1.5 / 3

✓ **- 1.5 pts The hypothesis are not based on the sample.**

5.2 interpretation (Corey Katz) **0 / 4**

✓ **- 4 pts** **Incorrect**

QUESTION 6

Which inference? (Chu Ching and Shannon Proctor) 25 pts

6.1 Ecog procedure (Chu Ching) **2 / 2**

✓ **+ 2 pts** **correct**

6.2 Ecog hypotheses (Shannon Proctor) **3 / 3**

✓ **- 0 pts** **Correct**

6.3 REM procedure (Chu Ching) **2 / 2**

✓ **+ 2 pts** **correct**

6.4 REM hypotheses (Shannon Proctor) **3 / 3**

✓ **- 0 pts** **Correct**

6.5 Bise procedure (Chu Ching) **2 / 2**

✓ **+ 2 pts** **correct**

6.6 Bise hypotheses (Shannon Proctor) **3 / 3**

✓ **- 0 pts** **Correct**

6.7 Bone marrow procedure (Chu Ching) **2 / 2**

✓ **+ 2 pts** **correct**

6.8 Bone marrow hypotheses (Shannon Proctor) **3 / 3**

✓ **- 0 pts** **Correct**

6.9 Screen time procedure (Chu Ching) **2 / 2**

✓ **+ 2 pts** **correct**

6.10 Screen time hypotheses (Shannon Proctor) **3 / 3**

✓ **- 0 pts** **Correct**

QUESTION 7

7 Margin of error (Chu Ching) **3 / 3**

✓ **+ 3 pts** **correct**

QUESTION 8

Delayed first bath (Shannon Proctor) 15 pts

8.1 hypotheses (Shannon Proctor) **0 / 3**

✓ **- 3 pts** **Incorrect**

This is incorrect and demonstrates a misunderstanding of this problem and the concepts involved. See the solutions.

8.2 calculations (Shannon Proctor) **3 / 3**

✓ **- 0 pts** **Correct**

8.3 which condition? (Chu Ching) **2 / 2**

✓ **+ 2 pts** **correct**

8.4 interpretation (Shannon Proctor) **4 / 4**

✓ **- 0 pts** **Correct**

8.5 recommendation (Baldi) **0 / 3**

✓ **- 3 pts** **Not directly relevant. Making a "recommendation" implies that causation is established, which goes beyond just finding a significant association.**

Stats 8 winter 2019 (Baldi) **Student UCI email:** bnchung@uci.edu
Final vB
Seat # L-9 **First name, Last name:** Brian, Chung

Keep your exam to yourself and your eyes on your work. No electronic devices are allowed at any time, except for a calculator. Cheating will be reported and result in a failed course.
Show your work for full credit. Use math symbols whenever appropriate. Use a significance level of 0.05 if necessary. *You have 2 hours.*

Problem 1 (22 points) < 30 min
A 2019 study aimed to compare the health impact of different physical activities in women with type-2 diabetes. At random, 10 women with type-2 diabetes participated in a resistance training program and 8 participated in an endurance training program. Body weight issues are always a concern for patients with diabetes. The participants' body mass index (BMI, in kg/m^2) was recorded at the beginning of the study and again after 10 weeks of training. The findings are show below (with the format mean ± standard deviation) as they appear in the published report.

	Before	**After**	**Change**
Endurance (n = 8)	28.50 ± 2.38	27.70 ± 2.20	–0.80 ± 0.67
Resistance (n = 10)	28.45 ± 3.32	27.56 ± 3.15	–0.89 ± 0.48

<u>Part 1</u>: What is the impact of <u>resistance</u> training for 10 weeks on BMI in this population?

a) Obtain a 95% confidence interval for the mean BMI change with resistance training. Show your full calculator work. $df = 10-1 = 9$

t-lower = invT (area: 2.5/100, df: 9) = -2.262157158 sto→A = A

t-upper = invT (area: 97.5/100, df: 9) = 2.262157158 sto→B = B

$\bar{x} \pm t^* \frac{s}{\sqrt{n}} \rightarrow \left(-0.89 + A \cdot \frac{0.48}{\sqrt{10}}, -0.89 + B \cdot \frac{0.48}{\sqrt{10}}\right) \rightarrow (-1.233, -0.547)$

b) The value of the margin of error for this confidence interval is: 0.343 ← $t^* \frac{s}{\sqrt{n}} = B \cdot \frac{0.48}{\sqrt{10}}$

c) Interpret your confidence interval in the context of this problem.
There is a 95% confidence level that the change in BMI due to resistance training in women is, on average, -1.233 to -0.547. If samples like this are repeatedly taken, then 95% of the samples will capture the true change in BMI of women doing resistance training. Therefore, there is a 95% probability that this interval (-1.233, -0.547) captures the true change in BMI in women due to resistance training.

d) Based on this confidence interval, can we conclude that resistance training for 10 weeks can help significantly reduce BMI in this population? Briefly explain your reasoning.
Yes we can. This confidence interval can be used for a null hypothesis test in which the null hypothesis is that there is no change in BMI, and so $\mu_0 = 0$. Because the interval does not capture μ_0, the null hypothesis can be rejected, and so resistance training, most likely and on average, changes BMI. In addition, because this interval only captures negative values, and because there's a 95% probability that this interval captures the true change, it is very likely that BMI decreases due to resistance training.

Part 2: Do the findings indicate that resistance and endurance training have significantly different effects on BMI change in the population of women with type-2 diabetes?

a) Write the corresponding null and alternative hypotheses.

H_0: $\mu_{endurance} = \mu_{resistance}$ H_a: $\mu_{endurance} \neq \mu_{resistance}$

b) Obtain the test statistic and *P*-value. Show your full calculator work.

endurance → 1 resistance → 2

2-samp T Test (Inpt: Stats, $\bar{x}1$: -0.8, $Sx1$: 0.67, n1: 8, $\bar{x}2$: -0.89, $Sx2$: 0.48, n2: 10, $\mu1$: ≠ $\mu2$, Pooled: No) →

$\boxed{\text{test-statistic: } t = 0.320 \quad \text{P-value} = 0.754}$

c) Conclude in the context of this problem.

Ho predicts that the probability of obtaining a test statistic at least as extreme as observed is 0.754. Because the probability that Ho predicts the observed results is so high (greater than 0.05), Ho cannot be rejected. Therefore, the hypothesis that there is no difference in BMI change between resistance and endurance training should not be rejected.

Problem 2 (9 points) < 12 min

Fluoroquinolone is a common antibiotic, but fetal risk is an important consideration when treating pregnant women. An observational study examined the medical records of 783 women exposed to a fluoroquinolone during the first trimester of their pregnancy and 3433 women with no such exposure during pregnancy.

Following are further details on the study. For each one, select the appropriate inference procedure:

A) One sample or matched pairs *t* for a mean
B) Two sample *t* for two means
C) ANOVA for several means
D) Chi-square goodness of fit
E) Chi-square for two-way table of counts

a) The researchers compared the percent of newborns with a birth defect in the two groups to see if there was a significant difference between the groups.
categorical variable

~~A~~ B C D E

b) The researchers examined the 783 women who had been exposed to a fluoroquinolone in their first trimester and sorted them according to the timing of the exposure (first month, second month, or third month). They compared the mean head circumference (in cm) of newborns in these three groups to see if it may be affected by the timing of fluoroquinolone exposure.

A B C D E

c) The researchers compared the mean maternal age in the two groups to verify that the two groups were not significantly different in this respect.

A B C D E

UCI email: bnchung @uci.edu

Problem 3 (11 points) < 15 min

Advertisement to children is tightly regulated, but social media content isn't. An Australia study published earlier this month examined whether social media influencers can influence children's food choices. The researchers created fake Instagram profiles for popular YouTubers, known in the media business as "influencers." At random, 176 Australian children (ages 9 to 11) were divided into three groups shown Instagram profiles that featured the influencer in different settings: with unhealthy snacks, with healthy snacks, or with non-food products. After viewing their assigned Instagram page, the children were offered a variety of snack options.

The researchers recorded how many calories of unhealthy food (in Cal) each child ate. Here are a summary of the findings and software output for the corresponding ANOVA *F* test:

Setting	N	Mean	St. Dev.
Unhealthy	59	384.83	141.21
Healthy	59	319.51	143.77
Non-food	58	292.24	146.85

```
Source of Variation            SS     DF            MS         F         P
Between Groups:         265532.1750     2   132766.0875    6.4076    0.0021
Within Groups:         3584585.0485   173    20720.1448
Total:                 3850117.2235
```

a) Write the corresponding null and alternative hypotheses.

H_0: $\mu_{unhealthy} = \mu_{healthy} = \mu_{non\text{-}food}$

H_a: the mean calorie intake of 1 or 2 settings is not equal to the other setting(s)

b) Assuming for now that the conditions for inference are met, interpret the findings in context.

The probability that Ho predicts results as extreme as observed is only 0.0021. Therefore, it is very unlikely that Ho would predict these results. Ho should be rejected in favor of Ha. Therefore, it is very unlikely that, on average, the mean calorie intake of unhealthy food is the same in all 3 settings. At least 1 of these settings differ in terms of calorie intake. In this case, the calorie intake appears to be highest in the unhealthy snacks setting, second highest in the healthy snacks setting, and lowest in the non-food products setting.

c) What are the conditions for this inference procedure?

A) One random sample AND normal data or large enough sample size.
B) Independent random samples AND normal data or large enough total sample size.
~~C) No conditions are needed for this inference procedure.~~
D) Normal data AND large enough ~~expected counts.~~
E) Independent random samples AND normal data or large enough total sample size AND similar sample standard deviations.
F) Random selection without repeated measures AND large enough expected counts.

d) Are the conditions for inference met?

A) No, none of them are met. **B)** No, only some are met. **C)** Yes, all conditions are met.

Problem 4 (7 points) < 10 min

In a genetic experiment, chickens with white feathers and a large head crest (WL) were bred with chickens with dark feathers and a small crest (DS). It was theorized that the genetic inheritance of these phenotypes may be based on two independent genes, from which a 9:3:3:1 ratio in second generation would be expected (with the phenotype WS being the most common and the phenotype DL the least common). The breeding program actually produced in second generation 111 WS chicken, 37 WL, 34 DS, and 8 DL.

Here is the software output for the corresponding chi-square for goodness of fit test.

Category	Observed	Historical Counts	Test Proportion	Expected	Contribution to Chi-Sq
WS	111	9	0.5625	106.875	0.15921
WL	37	3	0.1875	35.625	0.05307
DS	34	3	0.1875	35.625	0.07412
DL	8	1	0.0625	11.875	1.26447

N	DF	Chi-Sq	P-Value
190	3	1.55088	0.671

a) Write the corresponding null and alternative hypotheses.

Independent assortment is true

H_0: $WS = 190 \times \frac{9}{16} = 106.875$, $WL = 190 \times \frac{3}{16} = 35.625$

$DS = 190 \times \frac{3}{16} = 35.625$, $DL = 190 \times \frac{1}{16} = 11.875$

Independent assortment it's not true

H_a: $WS \neq 106.875$ or $WL \neq 35.625$ or $DS \neq 35.625$ or $DL \neq 11.875$ or a combination of these

b) Use the software output to conclude in context for the whole population of second generation chickens that would result from crossing WL and DS chickens.

A) The findings are consistent with an equal representation of the 4 phenotypes ($P = 0.671$), which does not imply that they are truly equally represented.

B) The findings are consistent with an equal representation of the 4 phenotypes ($P = 0.671$), indicating that they are truly equally represented.

C) There is strong evidence ($P = 0.671$) that the 4 phenotypes are not equally represented.

D) There is weak evidence ($P = 0.671$) that the 4 phenotypes are not equally represented.

E) The findings are consistent with the proportions expected under a genetic model with two independent genes ($P = 0.671$), which does not imply that they are truly inherited according to this model.

F) The findings are consistent with the proportions expected under a genetic model with two independent genes ($P = 0.671$), indicating that are truly inherited according to this model.

G) There is strong evidence ($P = 0.671$) that the 4 phenotypes are not represented according to the proportions expected under a genetic model with two independent genes.

H) There is weak evidence ($P = 0.671$) that the 4 phenotypes are not represented according to the proportions expected under a genetic model with two independent genes.

I) The test assumptions are not met and, therefore, we should not conclude anything.

UCI email: bnchung@uci.edu

Problem 5 (5 points each) < 25 min

For each following question, select the appropriate inference procedure. Then state the corresponding null and alternative hypotheses, using math symbols whenever possible.

a) The American Heart Association has formal recommendations for the maximum amount time that children should spend looking at screens, with a limit of 1 hour per day for children ages 2 to 5 years. A 2017 survey of parents in the United States found that the parents of children aged 2 to 5 reported that their children spend 2.6 hours per day in screen-based activities, on average.

A) One sample or matched pairs *t* for a mean
B) Two sample *t* for two means
C) ANOVA for several means
D) Chi-square goodness of fit
E) Chi-square for two-way table of counts

H_0: $\mu_{\text{screen time}} = 1$ hr/day

H_a: $\mu_{\text{screen time}} > 1$ hr/day

b) A 2019 study examined the effect of bone marrow transplants on the "physiological youth" of mice. The researchers compared random samples of young mice, old mice, old mice transplanted with bone marrow taken from young mice, and old mice transplanted with bone marrow taken from old mice. The mice were placed in transparent cages and videotaped to measure their activity level (in meters walked for the first 20 minutes after being placed in the cage).

A) One sample or matched pairs *t* for a mean
B) Two sample *t* for two means
C) ANOVA for several means
D) Chi-square goodness of fit
E) Chi-square for two-way table of counts

H_0: $\mu_{\text{young}} = \mu_{\text{old}} = \mu_{\text{old with young marrow}} = \mu_{\text{young with old marrow}}$

H_a: At least one of the mean activity level in one of these treatment groups is different from the other treatment groups.

c) Alzheimer's disease (AD) can be diagnosed using a detailed questionnaire resulting in a score on the Everday Cognition scale (Ecog, with higher scores indicating a worse cognitive impairment). This questionnaire is typically administered in person by a trained clinician, but a new online tool lets people self-administer the questionnaire. Out of concern that the two approaches may yield significantly different results, researchers compared a random sample of AD patients who took the questionnaire at a clinic and a random sample of patients who took the questionnaire online.

A) One sample or matched pairs *t* for a mean
B) Two sample *t* for two means
C) ANOVA for several means
D) Chi-square goodness of fit
E) Chi-square for two-way table of counts

H_0: $\mu_{\text{clinic}} = \mu_{\text{online}}$

H_a: $\mu_{\text{clinic}} \neq \mu_{\text{online}}$

d) French people commonly greet each other with a light kiss on each cheek ("la bise"). Is there a social preference for starting on the right versus the left cheek, or is it just a chance event? Researchers recorded a random sample of French individuals greeting each other with "la bise" and found the proportion who started with a kiss on the right cheek.

A) One sample or matched pairs *t* for a mean
B) Two sample *t* for two means
C) ANOVA for several means
D) Chi-square goodness of fit
E) Chi-square for two-way table of counts

H_0: $P_{\text{right cheek}} = 0.5$

$P_{\text{left cheek}} = 0.5$

H_a: $P_{\text{right cheek}} \neq 0.5$ and

$P_{\text{left cheek}} \neq 0.5$

e) We know that infants fall asleep easily when being gently rocked, but what would be the effect on adults? A 2019 study recruited a random sample of 18 healthy young adults to monitor their quality of sleep over two nights, measured by the time (in minutes) spent in REM sleep. In random order, on one night the participants slept in a stationary bed and on the other night they slept in a bed that gently rocks.

A) One sample or matched pairs *t* for a mean
B) Two sample *t* for two means
C) ANOVA for several means
D) Chi-square goodness of fit
E) Chi-square for two-way table of counts

H_0: $\mu_{\text{stationary bed} - \text{rocking bed}} = 0$

H_a: $\mu_{\text{stationary bed} - \text{rocking bed}} \neq 0$

Problem 6 (3 points) < 3 min

The margin of error of a confidence interval to estimate a population parameter accounts for the sample-to-sample variations in the value of the statistic

A) due to ~~nonresponse and response bias.~~
B) due to measurement errors and ~~response bias.~~
C) due only to measurement errors. ?
D) due only to the random sampling process.
E) due to anything that may influence the statistic as long as the population is Normal.
F) due to anything that may influence the statistic as long as the data are a random sample.
G) due to anything that may influence the statistic as long as the population is Normal and the data are a random sample.

Problem 7 (15 points) < 20 min

In the summer of 2016, the Cleveland Clinic started using delayed bathing for newborns (bathing no sooner than 12 hours after birth, "intervention"), hoping that the uninterrupted mother-infant time would help promote breastfeeding. But does it actually have any effect on breastfeeding? A study published earlier this month compared the rates of exclusive breastfeeding during the hospital stay in a random sample of 448 babies born at the hospital in the spring of 2016 who were bathed within 2 hours of birth ("preintervention") and a random sample of 548 babies born in the fall of 2016 who got a delayed bath ("postintervention"). The reported findings are shown below.

total $= 448 + 548 = 996$

Variable[a]	Total	Preintervention ($n = 448$)		Postintervention ($n = 548$)	
Exclusive breastfeeding, n (%)					
No	354 (35.5)	180 (40.2)	$\frac{448 \times 354}{996} = 159$ Sto → A	174 (31.8)	$\frac{354 \times 548}{996} = 195$ Sto → C
Yes	642 (64.5)	268 (59.8)	$\frac{448 \times 642}{996} = 289$ Sto → B	374 (68.2)	$\frac{642 \times 548}{996} = 353$ Sto → D

a) Write the null and alternative hypotheses corresponding to the study's stated objective.

No association b/w bath timing & breastfeeding

H_0: pre intervention, no breastfeed = 159
pre intervention, yes breastfeed = 289
post intervention, no breastfeed = 195
post intervention, yes breastfeed = 353

Association b/w bath timing & breast feeding

H_a: pre intervention, no breastfeed ≠ 159 or
pre intervention, yes breastfeed ≠ 289 or
post intervention, no breastfeed ≠ 195 or
post intervention, yes breastfeed ≠ 353 or any combination of these

b) Obtain the test statistic and *P*-value. Show your full calculator work.

$df = (2-1)(2-1) = 1 \cdot 1 = 1$

$$\chi^2 = \frac{(180-A)^2}{A} + \frac{(268-B)^2}{B} + \frac{(174-C)^2}{C} + \frac{(374-D)^2}{D} = 7.640107879 \quad \text{Sto} \to E$$

χ^2 cdf (lower: E, upper: 1E99, df: 1) → $\boxed{\text{P-value: } 0.0057 \quad \text{test-statistic: } \chi^2 = 7.640107879}$

UCI email: bnchung @uci.edu

c) What are the conditions for this inference procedure?

A) ~~One~~ random sample AND normal data or large enough sample size.
B) Independent random samples AND normal data or large enough total sample size.
C) ~~No conditions are needed~~ for this inference ~~procedure.~~
D) Normal data AND ~~large~~ enough expected counts.
E) Independent random samples AND normal data or large enough total sample size AND ~~similar sample standard~~ deviations.
F) Random selection without repeated measures AND large enough expected counts.

d) Conclude in the context of this problem.

Ho predicts a test statistic that is 7.64 or greater with a probability of only 0.0057. Therefore, it is very unlikely that Ho predicts the observed results, and so Ho should be rejected in favor of Ha. It is very unlikely that the proportions of babies who breastfeeds or not and who received a preintervention or postintervention bath closely follow the proportions predicted by Ho. Therefore, there is an association between the 2 tested categorical variables (whether or not a baby breastfeeds and the timing of the bath). It appears that babies, on average, are less likely to breastfeed when given a postintervention bath, as the actual proportion of babies who exclusively breastfed after a postintervention bath is lower than predicted.

e) Can we conclude from this study that delaying the first bath of newborns by at least 12 hours can be recommended as an effective way of increasing the rate of exclusive breastfeeding while the mother and infant are at the hospital? Briefly EXPLAIN your reasoning.

No. the proportion of babies who exclusively breastfeeds after a postintervention bath is lower than predicted while the proportion of babies who exclusively breastfeeds after a preintervention bath are higher than predicted. Therefore, on average, it appears that a postintervention bath is associated with a lower proportion of exclusive breastfeeding. Therefore, it is recommended that the hospital don't give newborns postintervention baths (baths that are at least 12 hours after birth) if they want to increase the rate of exclusive breastfeeding.