

STATS 8 w19 midterm1w19-vB

Brian Nhan Thien Chung

TOTAL POINTS

43 / 50

QUESTION 1

Usual Source of care 8 pts

1.1 number of variables 0 / 2

- **0 pts** Correct
- ✓ **- 2 pts** Incorrect

1.2 Interpretation 6 / 6

- ✓ **- 0 pts** Correct

- **3 pts** You are missing one correct answer
- **6 pts** You selected neither of the two correct answers
- **2 pts** You selected one wrong answer
- **4 pts** You selected two wrong answers
- **6 pts** Wrong

QUESTION 2

2 DBS 2 / 2

- ✓ **- 0 pts** Correct
- **2 pts** Incorrect(the correct answer is f)

QUESTION 3

3 Alcohol and Marijuana 6 / 6

- ✓ **- 0 pts** Correct
- **2 pts** Didn't put gender as variable
- **1 pts** Didn't put gender as categorical
- **3 pts** If put "percent morally acceptable alcohol and percent morally acceptable marijuana" as variables
 - **4 pts** Put variables as "alcohol morally wrong, alcohol morally right, marijuana morally wrong, and marijuana morally right"
 - **2 pts** Added "percent acceptable or percent wrong" as a variable.
 - **2 pts** Did not distinguish "opinion on morality of

alcohol" and "opinion on mortality of marijuana" as separate variables.

- **2 pts** Did not specify that response was "opinion on the morality of..."
- **4 pts** Incorrectly specified morality variables
- **2 pts** Missing or incorrect for one variable
- **2 pts** Put too many variables
- **6 pts** Incorrect
- **2 pts** Put quantitative instead of categorical

QUESTION 4

Climate change 13 pts

4.1 Slope interpretation 3 / 4

- **0 pts** Correct
- ✓ **- 1 pts** Did not say "on average"
- **1 pts** Not in context
- **1 pts** Missing units
- **4 pts** Not what asked
- **1 pts** Ambiguous
- **3 pts** This is not the correct interpretation of the slope in context
- **2 pts** incorrect unit increase
- **4 pts** This is not the correct interpretation of the slope
- **1 pts** Did not say "increase"
- **1 pts** Did not specify "...for every 1 degree Celsius increase in sea surface..."

4.2 y-intercept 2 / 2

- ✓ **- 0 pts** Correct
- **2 pts** Incorrect

4.3 r-squared 3 / 3

- ✓ **- 0 pts** Correct
- **2 pts** Incorrect

- **0.5 pts** Did not use percent

4.4 conclusions 4 / 4

✓ - **0 pts** Correct

- **2 pts** You are missing one correct answer
- **4 pts** You are missing two correct answers
- **2 pts** You selected one incorrect answer
- **4 pts** You selected two incorrect answers
- **4 pts** Wrong

QUESTION 5

Birth Data 13 pts

5.1 Histogram 4 / 4

✓ - **0 pts** correct

- **0 pts** Correct range
- **0 pts** correct distribution
- **2 pts** did not write bimodal(incorrect distribution) or missing distribution
- **1 pts** stretching range (If the range is too far from correct)
- **2 pts** wrong range (the range of values is incorrect)
- **1 pts** if it is bimodal it cannot be skewed or symmetric.
- **4 pts** Incorrect

5.2 days of the week 4 / 6

- **0 pts** Correct

✓ - **2 pts** Missing one correct answer

- **4 pts** Missing two correct answer
- **6 pts** no correct answers selected
- **2 pts** selected one wrong
- **4 pts** Selected two incorrect
- **6 pts** Incorrect

5.3 Comparing plots 3 / 3

✓ - **0 pts** Correct

- **1 pts** Mostly correct (if you had the right idea, but you did not complete the point)
- **2 pts** Mostly correct but too vague (on the right track, but not enough)

- **3 pts** wrong (far off topic) or Blank, we were not asking for a description of the two graphs.

QUESTION 6

Nurses 8 pts

6.1 Type of study 0 / 2

- **0 pts** Correct

✓ - **2 pts** Incorrect

6.2 Variables 3 / 3

✓ - **0 pts** Correct

- **1 pts** Be more specific (need more than death, from what?), or not proportion of deaths or mortality, or missing "mortality" or "death".

- **2 pts** the variable is death from heart disease and it is categorical (do not confuse the summary of the variable with the variable)

- **3 pts** Wrong Variable or Incorrect

- **2 pts** missing variable type

6.3 conclusion 3 / 3

✓ - **0 pts** Correct

- **3 pts** incorrect

Midterm1 vB

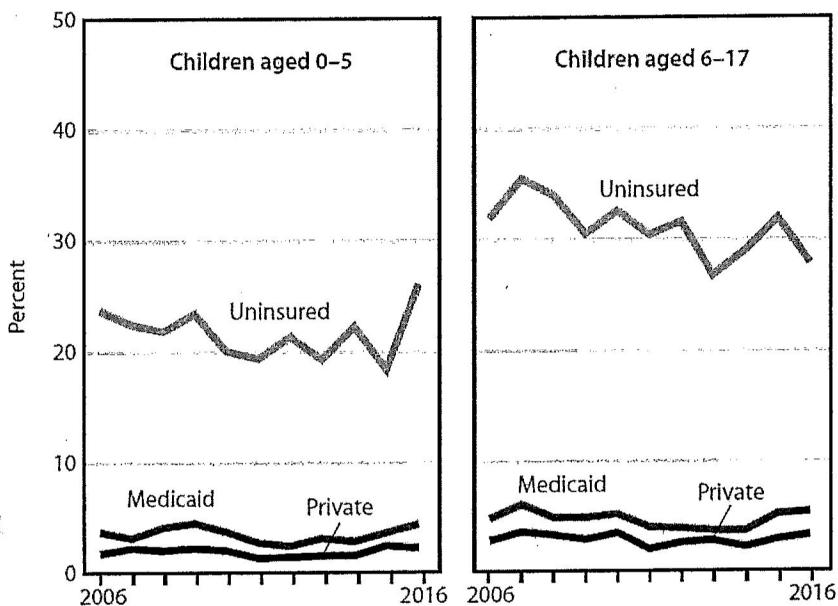
Seat # F3 Initials: BC Last name: Chung

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 an F course grade. You have ~ 45 minutes.

Problem 1 (< 7 min)

Children benefit from having a “usual source of care” (such as a designated doctor) for preventive health services and the treatment of medical conditions. So it is important for a nation to identify the children who have no usual source of care.

Figure 15. No usual source of care among children under age 18 years, by type of health insurance coverage: United States, 2006–2016



The National Health Interview Survey, conducted yearly in the United States, collected data on that topic. The findings were published in this figure.

[Note: “Medicaid” is a public health insurance coverage.]

Var 1: Health insurance coverage

Var 2: Usual source of care availability

- a) Not counting time (in years), how many variables are displayed in these time plots to describe the population of U.S. children? [Circle correct answer]

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

- b) Circle all correct interpretations of these graphs (there could be more than one).

Overall, during that period,

- A) there were more uninsured children aged 6–17 than uninsured children aged 0–5.
- B) uninsured children aged 6–17 were more likely to have no usual source of care than uninsured children aged 0–5.
- C) children aged 6–17 were more likely to be uninsured than children aged 0–5.
- D) in both age groups, rates of uninsured children showed cyclical variations whereas rates of children with private or Medicaid health insurance did not.
- E) in both age groups, uninsured children were more likely to have no usual source of care than children with private or Medicaid health insurance.
- F) in both age groups, children were more likely to be uninsured than to have private or Medicaid health insurance.

Problem 2 (< 3 min)

Deep brain stimulation (DBS) is emerging as a potential treatment for mood disorders unresponsive to conventional therapy. A 2018 study examined the effectiveness of orbitofrontal cortex DBS in a sample of 25 subjects suffering from depression who were implanted with intracranial electrodes to treat epileptic seizures. During the procedure, the subject's mood was quantified in real time with a composite mood score (a higher value indicating a more positive mood). The researchers compared the composite mood score for each patient when receiving, in random order, a real DBS (a very low and painless 6 mA current) and a sham DBS (no electric current). *represents baseline*

This study is

- A) a cross-sectional sample survey.
- B) an unscientific study using voluntary response.
- C) a completely randomized experiment.
- D) a case-control observational study.
- E) a longitudinal cohort study.
- F) a matched pairs/repeated measures experiment.
- G) a census (population data).

Problem 3 (< 3 min)

A 2018 Beliefs and Values survey by Gallup used a random sample of 1024 American adults to examine American's attitudes about the morality of alcohol and marijuana use. Respondents were asked, "Regardless of whether or not you think that it should be legal, for each one, please tell me whether you personally believe that in general it is morally acceptable or morally wrong."

Here are some findings displayed in the published report:

	Drinking Alcohol		Smoking Marijuana	
	% Acceptable	% Wrong	% Acceptable	% Wrong
U.S. adults	78	19	65	31
Gender				
Men	84	15	70	26
Women	73	23	59	36

List all the variables reported here, and specify whether each variable is quantitative or categorical.

attitude towards drinking alcohol: categorical

attitude towards smoking marijuana: categorical

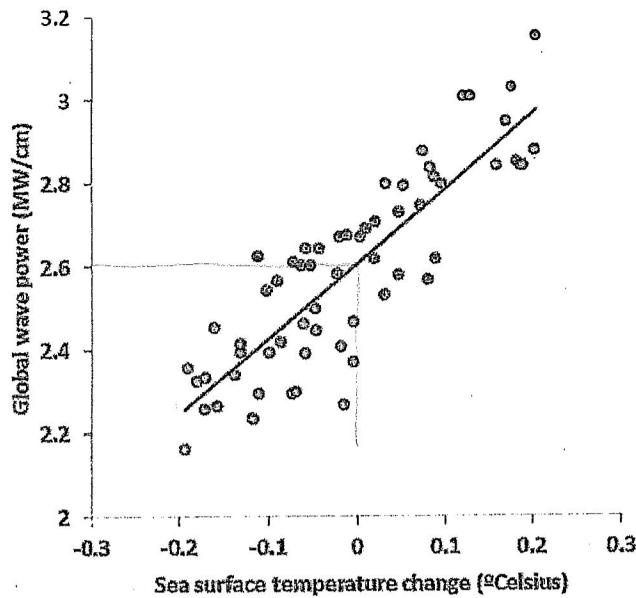
gender: categorical

Problem 4 (< 12 min)

Climate change is noticeable in increasingly warmer oceans globally. A 2019 study examined the relationship between sea surface temperature and how powerful waves are globally, by analyzing worldwide historical records for 1948 to 2017.

"Global wave power" reflected the power of waves all over the globe in a given year, and was measured in megawatts per centimeter (MW/cm). Ocean temperature in a given year was assessed as the overall "sea surface temperature change," measured in degrees Celsius relative to the average sea surface temperature in the late 20th century.

The findings are displayed in this graph. The line represents the least-squares regression line modeling global wave power as a function of sea surface temperature change. The numerical value of the slope is 1.804 and the linear correlation coefficient r is 0.861.



- a) Interpret the value of the slope in context.

For every 1°C increase in sea surface temperature, global wave power increases by approximately 1.804 mW/cm

- b) Rounded to one decimal place, what is the value of the y intercept? [Circle correct answer]

A) -0.3	B) 0.0
C) 2.1	D) 2.6

- c) What percent of the variation in global wave power is accounted for by sea surface temperature change in this regression model? $(0.861)^2 =$

74.1%

- d) Circle all appropriate conclusions from this study (there could be more than one).

- A) On average, globally, more powerful waves tend to be recorded for years with higher sea surface temperatures.
- B) Global wave power can be expected to exceed 4 MW/cm, on average, in years where the sea surface temperature would be a whole 1 degree Celsius higher than in late 20th century.
- C) The association between global wave power and sea surface temperature change is reasonably strong (clearly noticeable).
- D) Higher sea surface temperatures cause globally more powerful waves, on average.

Problem 5 (< 12 min)

The U.S. Social Security Administration keeps a record of all births in the United States. The following is an analysis based on birth records for the years 2000 to 2014.

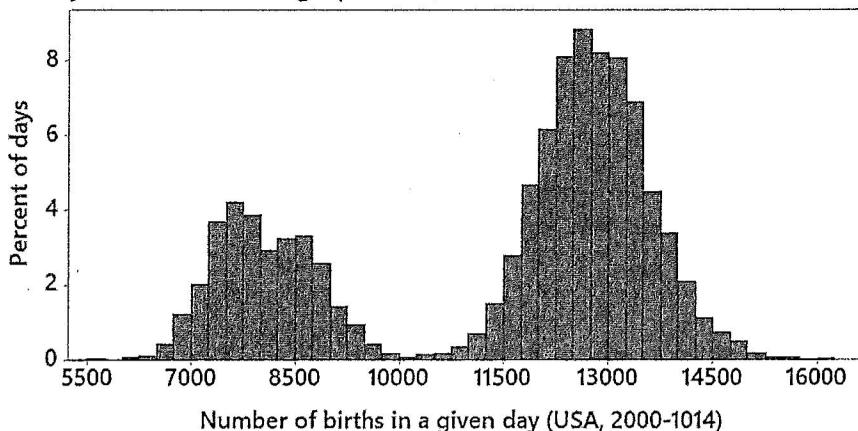
a) The data set is displayed graphically below. Use this graph to fill in the blanks.

For the 2000 to 2014 time period in the United States, the number of births in a day ranged approximately from

5250 to 16,250

births, and the distribution's shape is best described as

bimodal



b) The data set can be broken down by the day of the week that each birth occurred (Monday through Sunday), as shown here.

Day	Min	Q1	Median	Q3	Max
MON	6409	11616	12087	12506	14053
TUE	6325	12690	13198	13590	15645
WED	6620	12452	12932	13404	16081
THU	6471	12460	12958	13439	15590
FRI	6160	12179	12633	13170	15555
SAT	6159	8255	8574	8862	10499
SUN	5728	7259	7512	7779	8701

Circle all correct interpretations (there could be more than one).

In the United States, over the 2000–2014 period,

- A) The data set probably has typos or recording errors because the number of births cannot possibly be so much smaller on weekend days than non-weekend days.
- B) There is a clear association between number of births in a day and day of the week.
- C) Three-quarters of all Tuesdays had more than 13590 births.
- D) The minimum number of births was fairly similar for all seven days of the week, but the maximum varied noticeably across days of the week.
- E) The maximum number of births on weekend days was smaller than the first quartile for any of the non-weekend days.

c) Briefly describe how the second graph helps explain the pattern seen in the first graph.

In the second graph, during the week days, the number of births per day consistently range between 11,676 and 16,081. Because the week days consistently occupy this interval, they form the right-most peak in the first graph. Meanwhile, the number of births during the weekend consistently occupies a lower range, all below 10,499 births per day. As a result, they form the left-most peak on the first graph.

Problem 6 (< 8 min)

The Nurses' Health Study has been collecting health and behavioral data every two years since 1976 using the same very large, representative sample of female registered nurses in the United States. In 2004, a measure of optimism was included in the questionnaire filled by 70,021 study participants. A 2017 scientific article evaluated the association between optimism and mortality from heart disease in the 8 years following the optimism questionnaire. It reported that the proportion of women dying from heart disease was smaller for the group of women with the highest optimism level than for the group of women with the lowest optimism level. ^{2 cases}

a) This study is

- A) a cross-sectional sample survey.
- B) an unscientific study using voluntary response.
- C) a completely randomized experiment.
- D) a case-control observational study.
- E) a longitudinal cohort study.
- F) a matched pairs/repeated measures experiment.
- G) a census (population data).

b) Other than optimism level, what is the variable described in the 2017 study and is it quantitative or categorical?

Variable death from heart disease; quantitative /categorical

c) Can we conclude from the 2017 study that optimism helps prevent death from heart disease?

- A) No, because there were also heart disease deaths among women with the highest optimism.
- B) Not necessarily, because preexisting health issues are likely confounded with optimism level.
- C) Yes, because the study has a very large sample size.
- D) Yes, but only for the population of female registered nurses in the United States.

END

