Stats 7 winter 2015 (Baldi) – Midterm 1 version A SOLUTIONS

Problem 1 (10 points) < 12 min
Sixty-nine college freshmen participated in a 9-month paid prospective study. All participants were involved in a dating relationship at the beginning of the study. Every two weeks, participants answered a detailed questionnaire about their relationship and emotional wellbeing. Twenty-six of the participants experienced a breakup during the study. The average distress of these 26 participants at the time of breakup (0 weeks) and over the following 10 weeks was assessed from their questionnaire answers. Here are the findings:

Time since breakup (in weeks)	0	2	6	10
Average distress (unitless score)	3.3	3.1	2.9	2.6

(a) A scatterplot of the data indicates a linear relationship between time and distress. Provide the least squares regression line describing distress (y) as a function of time. Round the slope and intercept to 3 decimal places.
distress = 3.276 – 0.067 time (or y = 3.276 – 0.067 x)

(b) Explain briefly why you cannot use the regression line you computed to predict the average distress of a college freshman 6 months after a breakup (26 weeks).
This would be extrapolation (predication outside the range).

(c) If we measured time in days instead of weeks, which of the following would happen?

A) The slope and the correlation coefficient would both be different.
B) The slope and the correlation coefficient would both remain the same.
C) The correlation coefficient would be different but the slope would remain the same.
D) The slope would be different but the correlation coefficient would remain the same.

(d) The percent of variation in average distress explained by this regression model is **98.9%.**

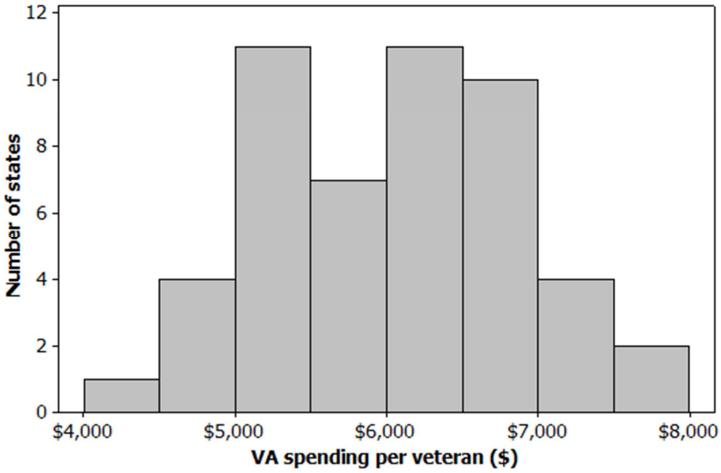
(e) This study is an example of

A) an observational study with one probability sample.
B) a cohort observational study.
C) a case-control observational study.
D) an experiment with a completely randomized or with a block design.
E) an experiment with a matched pairs or repeated measures design.
F) an unscientific study using a voluntary response sample.

Problem 2 (8 points) < 12 min
An NPR investigative report from earlier this month describes surprising state disparities in benefit spending for army veterans. Below are the per-veteran VA benefit spending (in dollars) for each of the 50 states in 2013.

4,480	4,587	4,797	4,935	4,999	5,019	5,056	5,082	5,085	5,261
5,409	5,414	5,437	5,465	5,474	5,493	5,508	5,715	5,715	5,793
5,860	5,885	5,925	6,186	6,190	6,192	6,207	6,237	6,257	6,287
6,302	6,340	6,438	6,440	6,531	6,620	6,676	6,677	6,694	6,719
6,780	6,816	6,859	6,890	7,044	7,172	7,295	7,462	7,626	7,689

(a) Make a histogram of these data using intervals of size \$500 starting at \$4,000 (\$4,000 to \$4,499.99; \$4,500 to \$4,999.99; etc.). Be sure to label your graph correctly.



(b) The median per-veteran VA benefit spending in 2013 was **\$6,191**.

(c) The distribution of state VA benefit spending in 2013 is best described as:

A) **roughly symmetric**
B) clearly bimodal
C) clearly right-skewed
D) clearly left-skewed

Problem 3 (2 points) < 3 min

A study funded by the World Wildlife Federation investigates snow monkeys. The researchers suspect that tail length may indicate if the snow monkey is an alpha male. A scatterplot of tail length and body length (both in inches) for 8 male snow monkeys revealed a linear association.

Here is software output for a linear regression analysis on these data. Based on this output, the correlation coefficient between tail and body length in male snow monkeys is:

r = **0.986** (√0.972)

```
Predictor       Coef  SE Coef      T      P
Constant      -7.369    1.085  -6.79  0.000
BodyLength   0.30122  0.02096  14.37  0.000

S = 0.198769   R-Sq = 97.2%
```

Problem 4 (10 points) < 5 min

A study tested the effectiveness of a bionic pancreas for glucose management among patients with type 1 diabetes. In random-order, each patient received therapy with a bionic pancreas for 5 days and therapy with their own insulin pump for 5 days. Researchers recorded for each patient in each condition the mean plasma glucose level (obtained every 2 hours) and the percentage of time that the patient had a low glucose level (<70 mg/dl). They found that among adolescent patients, the mean plasma glucose level was lower on average during the bionic-pancreas period than during the control period (138 vs. 157 mg/dl), but the percentage of time with a low plasma glucose reading was similar on average during the two periods (6.1% and 7.6%, respectively).

(a) This study is an example of

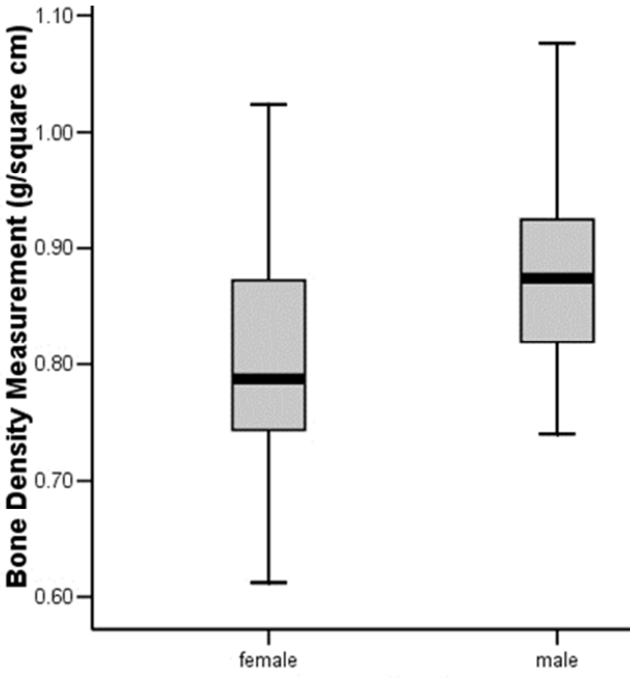
A) an observational study with one probability sample.
B) a cohort observational study.
C) a case-control observational study.
D) an experiment with a completely randomized or with a block design.
E) an experiment with a matched pairs or repeated measures design.
F) an unscientific study using a voluntary response sample.

(b) What are the two response variables studied and are they quantitative or categorical?

Variable 1: **plasma glucose level, quantitative** / ~~categorical~~

Variable 2: **percentage of time with low plasma glucose, quantitative** / ~~categorical~~

Problem 5 (6 points) < 5 min
A clinical trial of an osteoporosis drug enrolled about 600 adult subjects diagnosed with osteoporosis. The graph on the right shows the distribution of bone densities after treatment, for each gender.



(a) The interquartile range (IQR) for **men** is roughly **0.1 g/cm^2**.

(b) **50%** of **men** have a bone density lower than 0.88 g/cm^2.

(c) 75% of **women** have bone density above roughly **0.75 g/cm^2**.

Problem 6 (6 points) < 5 min
Acetaminophen is a "pain killer" medication commonly used to reduce physical pain. Could it also help with hurt feelings? A study recruited 62 healthy undergrads. By random assignment, half of the participants ingested a daily dose of 1,000 mg of acetaminophen, and the other half took the same dose of placebo. Each evening, participants used the Hurt Feelings Scale to report how much social pain they had experienced that day. The study found that participants' hurt feelings had decreased significantly by the end of the study in participants who took acetaminophen but not in participants who took the placebo.

(a) This is an example of

A) an observational study with one probability sample.
B) a cohort observational study.
C) a case-control observational study.
D) an experiment with a completely randomized or with a block design.
E) an experiment with a matched pairs or repeated measures design.
F) an unscientific study using a voluntary response sample.

(b) What is the response variable studied and is it quantitative or categorical?

Variable: **social pain/hurt feelings (Hurt Feeling Scale) quantitative** / ~~categorical~~

Problem 7
(2 points)
< 3 min

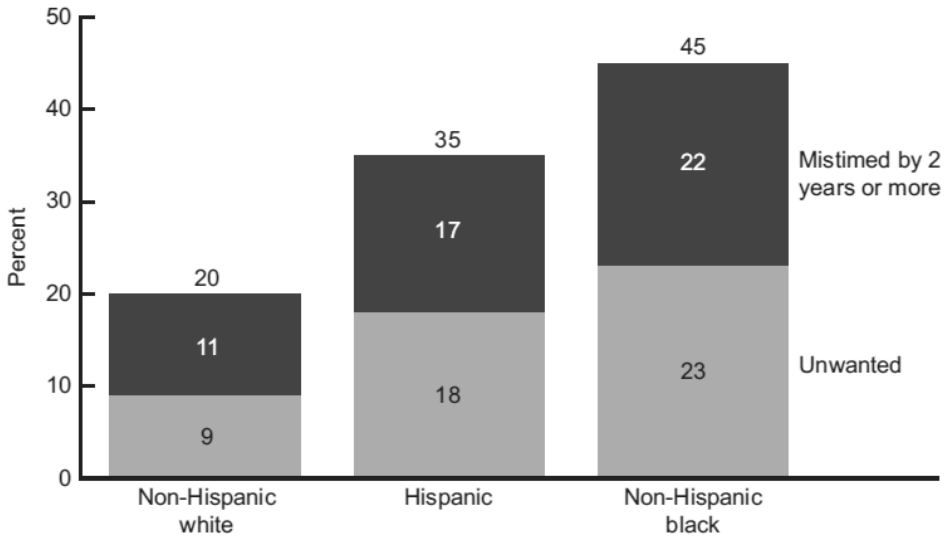


Figure 4. Percentage of births that were unwanted or mistimed by 2 years or more at conception, by Hispanic origin and race of mother: United States, 2006–2010

The latest National Health Statistics Reports on births in the United States reveals how "the intendedness of births varies considerably by the race and ethnicity of the mother." Figure 4 (above) shows the percentage of births in 2006–2010 that were either unwanted or mistimed by 2 years or more, broken down by ethnicity of the mother. This graph is a

A) histogram.
B) boxplot.
C) bar graph that could be made into one pie chart.
D) bar graph that cannot be made into one pie chart.