# 2013A IS4800/CS6350 Midterm Exam
## Closed Book, Closed Notes, 100 mins max

Name _____

**When asked to specify a "Study Design" please refer to the following list:**
Ethnographic, Descriptive, Correlational, Demonstrative, Between-Subjects Experimental
**You may modify these with the following prefixes:**
Univariate, multivariate, N-factor, N-level (for integer N).

**When asked what statistics you would use, be as specific as possible, e.g., "Mean and standard deviation of nose-picking frequency" or "t-test for independent means on scream intensity with torture method as the independent variable."**

**When asked "What kind of test is this?", choose from the following list:**
t-test for independent means, Chi-square goodness-of-fit, Chi-square test for independence, Pearson correlation

**When asked for a "Research Model", draw a boxes and arrows diagram depicting variables and their relationships. Label the boxes with the role of the variable(s) contained (IV, DV, etc.) as well as the name(s) of the variable(s).**

**When asked to "Interpret the results" of a test, you should write the results in both English and publication format. Example: "There were no significant differences in performance between the Jacuzzi and Sauna groups, t(42)=5.67, n.s."**

1. **(10%)** For each of the following measure, indicate (via checks) the kind of descriptive statistics you would use. Assume interval and ratio measures are approximately normal unless noted.

|  | Mean | Median | Mode | StdDev | Inter-quartile range |
|---|---|---|---|---|---|
| Q1. Latte Size (Sm,Med,Lg) |  |  |  |  |  |
| Q2. Servings of fruit per day someone eats. |  |  |  |  |  |
| Q3. Score on a composite scale (8 items, 7-pt ea) |  |  |  |  |  |
| Q4. Height, negative skew. |  |  |  |  |  |
| Q5. Age, with outliers. |  |  |  |  |  |
| Q6. Smokes (yes/no) |  |  |  |  |  |

2. You are the new manager of customer support for Exposed Networks, a cell phone provider specializing in nudist customers. The two primary performance metrics for your department are customer satisfaction and cases resolved per day. You are considering moving from voice-based customer support to a purely SMS- (texting)-based service.

**2a. (10%)** Design a questionnaire to assess satisfaction. It should include a three-item, one-factor scale (composite measure) that uses Likert items, including one negated item.

**(this page and next: 10%)**

2b. Go back and fill out your questionnaire (pretend you are a customer). What is your composite score for satisfaction?

_____

2c. To assess the quality of your questionnaire you give it to 30 nudists who have been long-time customers, wait a week, then give them the questionnaire again. What kind of quality assessment of your composite measure is this (name)?

_____

2d. What statistical test would you use on the data in 2c to demonstrate quality?

_____

2e. You now check to see if your satisfaction measure is related to the number of months each nudist has been a customer. What kind of quality assessment of your measure is this (name)?

_____

You have a small sample of customers take your survey and your statistician gives you the following visualization of the results.
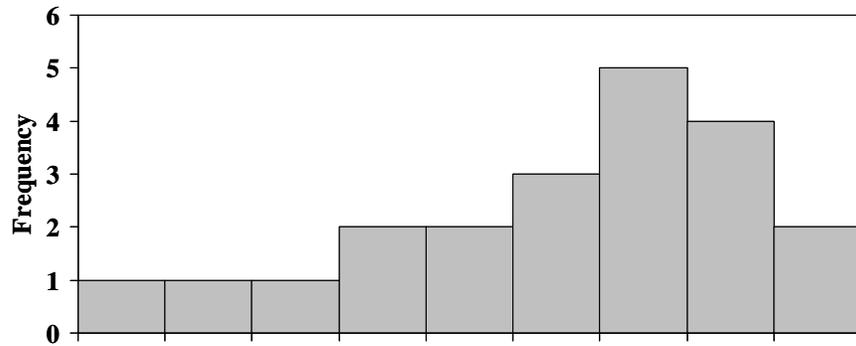


2g. What kind of a graph is this? _____

2h. Describe the distribution: _____

**2i. (20%)** Sketch a proposal for an experimental study using measures and manipulations described on the bottom of pg 2.

2i, continued

2i, continued

**3. (10%)** For each of the following, indicate the type of the study (experimental, correlational, descriptive, or demonstration).

3a. Exposed Networks knows their customers have issues carrying their phones, so they develop a suction cup-affixed holster, give it to 30 customers and, after a month, ask them how much they would be willing to pay for it.

_____

3b. Exposed wants to evaluate the impact of raising their rates on satisfaction, so they pick 100 customers at random and raise their rates and pick another 100 customers at random to compare them to (without raising rates), evaluating satisfaction in each group after 30 days.

_____

3c. Exposed measures customer age and number of calls placed per month for a sample of customers to determine if they should offer an age discount.

_____

3d. Exposed has customer service ask all of the customers who call in whether they belong to AARP or not.

_____

3e. You pick a Starbuck's and measure the number of people in line and wait time for service to see if you can predict how long it will take to get a cup of coffee.

_____

3f. In which of the above studies (3a-e) can you infer causality?

_____

**4. (10%)** Exposed Networks wants to get into the nudist dating market, so they add a FaceBook-style section to their website for all customers. They recruit a group of nudists, and randomly assign half of them to receive a link to the new web feature for free, and wait-list-control the other half. Data collected after 3 months is shown. Describe all appropriate statistical analyses enumerated in class and perform those you can.

| Nudist | Group | Satisfaction |
|--------|---------|--------------|
| Joe | Web | 7 |
| Sally | Control | 6 |
| Bill | Control | 3 |
| Martha | Web | 8 |
| Sue | Web | 9 |
| Mary | Control | 4 |

**5. (10%) Critique the following study design using the items below by putting letters next to the related criticisms.**

Exposed Networks marketing proposes to do a univariate, one factor, two-level between-subjects experiment to evaluate the impact of large phone keypads on text messaging speed while naked.  They approach <u>4 people walking down the street in front of their downtown office (a)</u> and <u>tell them that someone has been hurt in an accident and they need them to text message for an ambulance (b).</u>  <u>The first two subjects (approached in the morning) are given the large keypad phone and the second two subjects (afternoon) are given a normal phone.(c)</u>  <u>Before they start, three video cameras are pointed at them, heartrate and blood pressure sensors attached to their arm, and a motion-sensing cyberglove put on their dominant (texting) hand.(d)</u>  <u>The researchers find the significance of the t-test as p=0.34, and conclude that keyboard size has no effect on texting speed.(e)</u>

____ lacks control of extraneous variable

____ incorrect inference

____ biased sample

____ lacks ecological validity

____ does not address "respect for persons"

**6. (10%) Concepts**

6a. What is a Type I error?

_____

_____

6b. What is the "p value" (aka, "level of significance") for a statistical test?

_____

_____

6c. What kind of sampling method do most statistics (covered in class) assume you are using?

_____

_____

6d. What is the internal validity of a study?

_____

_____

6e. Describe one aspect of an experimental design that can affect the internal validity of a study.

_____

_____

6f. What is $\beta$?

_____

_____

## 7. (10%) R

7a. Describe this result in English and publication format.

```
> cor.test(c$Age,c$AvgCallCost)

Πεαρσονϵσ προδυχτ–μομεντ χορρελατιον

data: c$Age and c$AvgCallCost
t = -14.7207, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval: -0.9164725 -0.7985614
sample estimates: cor -0.8693912
```

_____

_____

_____

_____


7b. Describe this result in English and publication format.

**Phone**

| | Observed N | Expected N | Residual |
|---|---|---|---|
| Nokia | 8 | 4.3 | 3.7 |
| Motorola | 3 | 4.3 | -1.3 |
| iPhone | 2 | 4.3 | -2.3 |
| Total | 13 | | |

```
> chisq.test(c(8,3,2))

Χηι–σϑυαρεδ τεστ φορ γιϖεν προβαβιλιτιεσ

data:  c(8, 3, 2)
X-squared = 4.7692, df = ?, p-value = 0.09212
```

_____

_____

_____

_____

_____

_____