FINAL EXAM, NACP 204, **Research Methods**. Answer all 10 questions. Note that some offer you the choice of two or more options. For some questions you will need the abbreviated Chi-Square table copied at the end of this paper. (60 marks total)

1. As an Agricultural Extension agent you have become concerned by the extremely high incidence of Bilharzia (a fluke disease) among workers in a lowland region of your province. Somehow you obtain a random sample of 60 workers and cross-classify them according to blood type and infection. This cross-classification is reported in the table below:

|  | Blood Type | | |
|---|---|---|---|
|  | O | A | B |
| Bilharzia YES | 18 | 5 | 7 |
| Bilharzia NO | 6 | 19 | 5 |

What is $H_0$? (3 marks)
What is the Chi-Square value for this study? (Please show your work.) (4 marks)
From the Chi-Square table, one learns that P is between _____ and _____. (2 marks)
What do you decide about the null hypothesis? (1 mark)

2. In an experimental harvest of a large and intensively studied crocodile population, 25 experienced hunters will be licensed to kill up to 5 crocodiles each. Closely monitored, the harvest will be terminated as soon as 80 crocs have been killed. Because all crocs have value in the skin/meat market, every hunter will want to get his 5 before the hunt ends—so maybe he should kill the first 5 he encounters. But hides & meat from larger animals are more valuable than from smaller animals, so maybe the best strategy would be to hunt selectively for the big crocs. For this problem, you may assume that the encounter-probability is the same for every crocodile in the very, very large study-wetland. Here are some data.

| Size Class (m) → | 1-2 | 2-3 | 3-4 | >4 | Total |
|---|---|---|---|---|---|
| % of wild crocs in size class | 40% | | 20% | 10% | 100% |
| Expected kill under non-selectivity hypothesis | | | | | 80 |
| Observed kill by actual hunters | | 25 | 25 | 20 | 80 |

(a) Fill the blanks in the incomplete datasheet. (3 marks)
(b) Which of the following hypotheses ($H_1$ or $H_2$) is the appropriate $H_0$ for a Chi-Square test?
    $H_1$: Hunters will harvest crocodiles in proportion to their relative abundance.
    $H_2$: Hunters will selectively choose larger crocodiles. (2 marks)
(c) What is the critical value of Chi Square for an $\alpha$-level of $P = 0.05$? (1 mark)

3. **There are two options for Question 3. Select and answer only one.**

**OPTION A.** Explain concisely the relative advantages of **fully structured interviews** *versus* **loosely structured interviews**. Indicate the conditions (and associated research questions) under which one interview-strategy would be superior to the other. (8 marks)

**OPTION B.** A large seed-company is seeking to produce a cultivar of wheat that can be gainfully planted under a wide variety of conditions; the company has four candidate-varieties, which I'll call A, B, C, and D. As a very preliminary experiment, you decide to test the candidate-varieties under all possible combinations of four fertilizer mixes and four irrigation regimes. (You'd like to have lots of reps, but for a first test, well….) A schematic of your research-design looks like this:

|  |  | Irrigation Regimes | | | |
|---|---|---|---|---|---|
|  |  | Regime 1 | Regime 2 | Regime 3 | Regime 4 |
|  | Fert. Mix 1 | A |  |  |  |
| Fertilizer | Fert. Mix 2 |  |  |  |  |
| mixes | Fert. Mix 3 |  |  |  |  |
|  | Fert. Mix 4 |  |  |  |  |

(a) In the design-schematic above, indicate by letter an appropriate deployment of the candidate-varieties across treatment-combination cells; I've entered one letter to get you started. (3 marks)

(b) Write the model statement for your arrangement; again, I'll get you started. (3 marks)

YIELD = f(CANDIDATE-VARIETY, IRRIGATION-REGIME,…

(c) What, if anything, can you say about the interaction between two or more of the experimental variables? Explain if necessary. (2 marks)

4. Below is a set of alternative hypotheses that led to t-tests. For each $H_A$, say whether the test should be one-tailed or two-tailed.

(a) Antibiotics kept at 5ºC last longer than antibiotics kept at 20ºC. (1 mark)
(b) On average, Breed A hens and Breed B hens lay different numbers of eggs. (1 mark)
(c) The treatment will result in improvement. (1 mark)
(d) I think that **average for Population A ≠ average for Population B**. (1 mark)

5. **There are two options for Question 5. Select and answer only one.**

**OPTION A.** A researcher is testing the effect of two different diets (Diet A and Diet B) on weights (kg) of broilers raised for 4 weeks. The researcher has an 8-bird experimental grow-out pen in each of two very different locations on his University campus. It's not easy, but 4 birds in each location are fed exclusively on Diet A and four are fed exclusively on Diet B. The data look like this:

|  | Diet A | Diet A | Diet A | Diet A | Diet B | Diet B | Diet B | Diet B |
|---|---|---|---|---|---|---|---|---|
| Location 1 | 1.4 | 1.5 | 1.5 | 1.6 | 1.3 | 1.4 | 1.4 | 1.4 |
| Location 2 | 1.3 | 1.2 | 1.2 | 1.3 | 1.2 | 1.2 | 1.1 | 1.4 |

The researcher decides to analyze her data with fully factorial ANOVA using the following model: **Weight = f(Diet, Location, Diet*Location)**.

(a) Briefly explain the meaning of the interaction term. (5 marks)

The analysis produces the following table:

|  | SS | df | MS | F | p-value |
|---|---|---|---|---|---|
| Rows | 0.225625 | 1 | 0.225625 | 57 | 6.77E-06 |
| Columns | 0.050625 | 1 | 0.050625 | 12.78947 | 0.003809 |
| Inter | 0.000625 | 1 | 0.000625 | 0.157895 | 0.698079 |
| Within | 0.0475 | 12 | 0.003958 |  |  |
| Total | 0.324375 | 15 | 0.021625 |  |  |

(b) What is the row-variable? (1 mark)
(c) Is the row-variable statistically significant? (1 mark)
(d) What is the column-variable? (1 mark)
(e) Is the column-variable statistically significant? (1 mark)
(f) Comment appropriately on the interaction term. (1 mark)


**OPTION B.** A large and extremely wealthy UM Church in Dallas, Texas, USA, has established a program "to do something about food security in Africa." The founders of the program heard a missionary talk about a small village in southern Masvingo Province, Zimbabwe. Knowing that you are an expert in agricultural research, these people are willing to give you $100,000US to help the people of that village avoid food shortages, particularly during droughts. You are absolutely determined to spend the money wisely, but you know only these few facts about the village: (1) its resident population is about 500; (2) many men work in Harare or even South Africa, so > 50% of the farmers are women; (3) the rainfall for 2007 was 675mm; (4) outside food assistance (you don't know what) was needed in 2022. You also heard a rumor that in some years baboons raid the maize fields, but you don't know whether that's true.

So, concisely explain what information you would need most in order to spend the money wisely, and explain how you would use techniques covered in this course to obtain that information. (10 marks)


6. Designate by letter which scatterplot is most accurately described by the r-coefficients below. I'll do one for you. (5 marks)

A     B     C     D     E     F
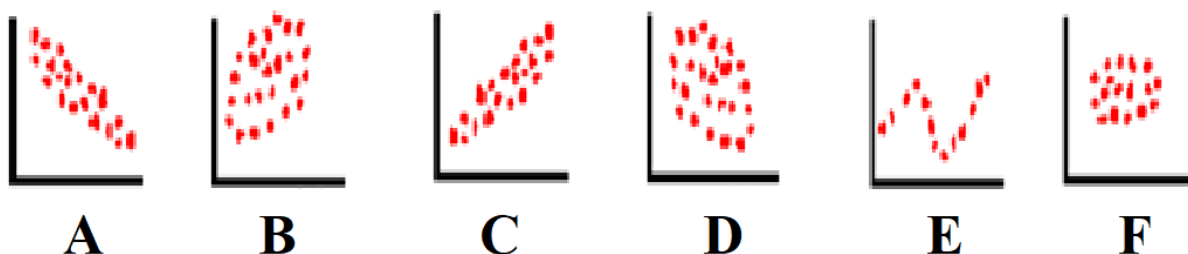
**Choices for r:**

**r ≈ +0.9   r ≈ -0.9   r ≈ + 0.3   r ≈ -0.3   r ≈ 0   This scatterplot is curvy; don't calculate r.**

(a) The best description for scatterplot A is _____ .          (1 mark)
(b) The best description for scatterplot B is _____ .          (1 mark)
(c) The best description for scatterplot C is _____ .          (1 mark)
(d) The best description for scatterplot D is _____ .          (1 mark)
(e) The best description for scatterplot E is _____ .          (1 mark)
(f) The best description for scatterplot F is **r = 0**.          (done)

7. A dairy scientist is interested in the affect of altitude on milk production in Zimbabwe. She discovers 10 very well-run dairy operations with genetically similar Holstein cows. For each dairy she learns the altitude (in meters above mean sea level, or AMSL) and the average per-cow daily milk production (in liters). Those data are summarized in the table below.

| Altitude (m AMSL) | 780 | 847 | 922 | 963 | 981 | 1005 | 1018 | 1037 | 1098 | 1150 |
|---|---|---|---|---|---|---|---|---|---|---|
| Production (liters/cow-day) | 27.2 | 29.4 | 28.8 | 27.9 | 17.0 | 22.6 | 29.1 | 20.5 | 20.1 | 21.9 |

The scientist fits a regression (least-squares) line to the data. The computer output is summarized in the table below.

| | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% |
|---|---|---|---|---|---|---|
| Intercept | 47.44076699 | 11.7382988 | 4.041537 | 0.003728 | 20.37220143 | 74.50933256 |
| X (m AMSL) | -0.023457573 | 0.011908959 | -1.96974 | 0.084377 | -0.050919682 | 0.004004537 |

(a) Write the regression (least-squares) equation; round coefficients to 2 decimal places.    (3 marks)
(b) What is the P-value for the slope coefficient (round to 2 decimal places)?    (1 mark)
(c) Which coefficient—slope or intercept—is more important for evaluating the effect of altitude on milk production?    (1 mark)
(d) What is the best estimate for the number of liters milk/cow-day that would be produced cows maintained at 1000mAMSL?    (3 marks)

8. Consider the following question, which might appear on a questionnaire examining attitudes towards a proposed governmental policy. (You may assume that questionnaire recipients know what "Proposal X" is. The question is in **bold italics** below.)

***"Do you agree that implementation of Proposal X would probably be cost-ineffective? (YES or NO _____ )"***

What is the most serious problem with the question? How would you rewrite it to make it better? (4 marks)

9. **There are three options for Question 9. Select and answer only one.**

**OPTION A**. If a researcher intends to perform statistical inference, what basic kind of sample does she or he need? (2 marks)

**OPTION B**. In the class you took this semester, what did your instructor claim was the most important thing to remember about writing a <u>solicited</u> proposal? (2 marks)

**OPTION C**. What is the basic difference between the Internet search engines ***scholar.google.com*** and ***google.com***? (2 marks)

10. **There are two options for Question 10. Select and answer only one.**

**OPTION A.** A researcher sent questionnaires to a random sample of 100 AU graduates. Fifty of the questionnaires were returned, and 80% of the respondents reported that they were "Very well satisfied" with their University experience. Would you be willing to generalize from this study to the entire population of graduates? Why or why not? (3 marks)

**OPTION B**. Explain the relative advantages and disadvantages of audio-recording interviews versus taking interview-notes by hand. (3 marks)
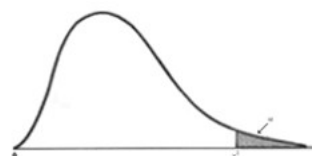
# Chi Square Distribution Table

| d.f. | $\chi^2_{.25}$ | $\chi^2_{.10}$ | $\chi^2_{.05}$ | $\chi^2_{.025}$ | $\chi^2_{.010}$ | $\chi^2_{.005}$ | $\chi^2_{.001}$ |
|------|------|------|------|------|------|------|------|
| 1 | 1.32 | 2.71 | 3.84 | 5.02 | 6.63 | 7.88 | 10.8 |
| 2 | 2.77 | 4.61 | 5.99 | 7.38 | 9.21 | 10.6 | 13.8 |
| 3 | 4.11 | 6.25 | 7.81 | 9.35 | 11.3 | 12.8 | 16.3 |
| 4 | 5.39 | 7.78 | 9.49 | 11.1 | 13.3 | 14.9 | 18.5 |
| 5 | 6.63 | 9.24 | 11.1 | 12.8 | 15.1 | 16.7 | 20.5 |
| 6 | 7.84 | 10.6 | 12.6 | 14.4 | 16.8 | 18.5 | 22.5 |
| 7 | 9.04 | 12.0 | 14.1 | 16.0 | 18.5 | 20.3 | 24.3 |
| 8 | 10.2 | 13.4 | 15.5 | 17.5 | 20.1 | 22.0 | 26.1 |