UNL Statistics PhD Qualifying Exam - January 2025

Print Your Qualifying Exam ID: _____

1. **8:00 am – 1:00 pm**

2. **Open book, open note. Bring your resources with you. You will not be allowed to go back up to your office to get additional notes. You may not use your phones for any reason.**

3. **Bring a calculator.**

4. **You may bring a flash drive with any SAS and R programs that you may possibly want to use.**

5. **You will have access to software available on the computers in the lab. You may NOT bring your own computer. You will NOT have access to the internet. If there is an `R` module that you need, our IT person will download it for you.**

    - **Answers to each problem should be saved as a single document in a directory on the USB drive you were provided.**
    - **Any programs and output for a question should be included in the directory for that problem.**

Day 2

1. (100 points) Let $e_1, \ldots, e_n$ be an independent random sample of size n from t_ν distribution and $Y_i = \mu + e_i$, here the true value of μ is 3. Two statistics for estimating μ are the sample mean and median.
   (a) (40 points) Design a simulation study (using $1,000$ repeats) to estimate the bias, variance, and mean squared error for both statistics. Your code must be enclosed as a function, otherwise 15 points will be deducted.
   (b) (30 points) For $n = 5, 15$, and 50 and $\nu = 2, 5$, and 30 estimate the bias, variance, and mean squared error for both statistics. Then use a grouped boxplot to plot the estimated **bias** for both statistics (mean v.s. median) for the case $n = 50, \nu = 5$, the plot needs to be well labeled.
   (c) (30 points) Write up a summary of your results including recommendations about when each statistic should be used. Interpretation is needed.

2. (100 points) Several bird behavior researchers set up an experiment to determine how house sparrows (hosp), a species of bird, reacted to four treatments assigned to four bird feeders over four periods. The treatments were designed to repel house sparrows from landing on a bird feeder. The four treatments were two line treatments: Hoop with 20C1 line and hoop with 20C2 line (20C1, 20C2), one Hoop only treatment(hoop) and one un-treated control (Control). During a period, each feeder was assigned a different treatment and by the end of the experiment, all four feeders had received each treatment. Data values were the total bird counts on a feeder for a 3-day period. File Jan25_open_book_1.sas contains the data and the program the researcher used for the following report. The intent of the researchers was to see if the mean bird counts for combined treatments 20C1+20C2 differed from combined Hoop+control treatments, if the two-line treatments differed (20C1 vs 20C2) and if Hoop differed from control.

   The researchers' report follows.....

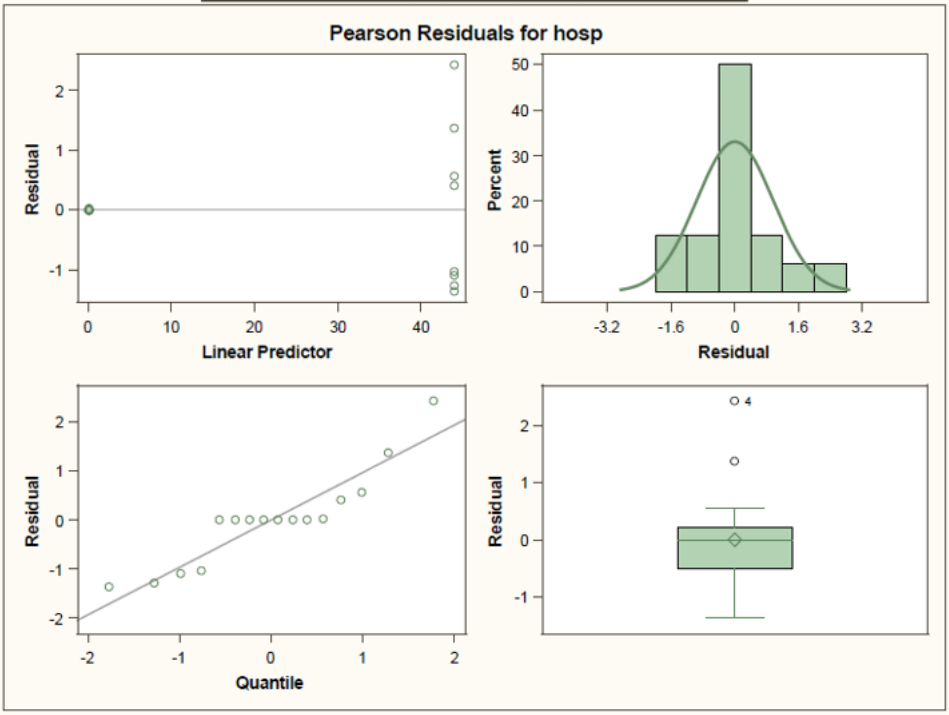
   The experiment was a completely randomized design with a single treatment factor (line) with two levels - presence and absence of line. The analysis produced the following results:

Fit Statistics	
-2 Res Log Likelihood	141.04
AIC (smaller is better)	147.04
AICC (smaller is better)	149.44
BIC (smaller is better)	148.95
CAIC (smaller is better)	151.95
HQIC (smaller is better)	146.86
Pearson Chi-Square	14444.88
Pearson Chi-Square / DF	1031.78

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
line	1	14	7.46	0.0162

line Least Squares Means

line	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
0	44.0000	11.3566	14	3.87	0.0017	44.0000	11.3566
1	0.1250	11.3566	14	0.01	0.9914	0.1250	11.3566



These results show that the only real effect of the experiment was presence/absence of lines over a feeder and that we should proceed with applying for a patent of the hoop with lines as an effective way of deterring house sparrows from bird feeders.

1. A review team led by the Statistics Department Comprehensive Exam Committee recommended that the report be rejected because the statistical analysis was flawed. Briefly explain in what ways the analysis was flawed.
2. Write a paragraph to correctly describe the experiment and hence the appropriate model to use for the analysis.
3. Give the program statements you would use for a corrected analysis.
4. Write a short report containing your corrected conclusions and recommendations. Please include plots and/or relevant output from your analysis.

```
Data sp;
input loc$ per trt$ line hosp; datalines;
A 1 20C2 1 1
A 2 20C1 1 0
A 3 HOOP 0 3
A 4 CONT 0 122
B 1 20C1 1 0
B 2 HOOP 0 0
B 3 CONT 0 62
B 4 20C2 1 0
C 1 CONT 0 57
C 2 20C2 1 0
C 3 20C1 1 0
C 4 HOOP 0 9
D 1 HOOP 0 11
D 2 CONT 0 88
D 3 20C2 1 0
D 4 20C1 1 0
proc print; run;
proc glimmix data=sp plots=pearsonpanel;
class loc per trt line; * 20C1 20C2 CONT HOOP;
 model hosp = line ;
lsmeans line ; run;
```