MATH 476 Statistics

Fred J. Hickernell | Final Exam | Monday, May 4, 2026

Instructions:

i. *This test has **FOUR** question(s). Attempt all. The maximum number of points is **100**.*

ii. *The time allowed is 120 minutes.*

iii. *This test is closed book, but you may use **8** double-sided letter-size sheets of notes.*

iv. *No calculators or other electronic devices are allowed. Phones must be placed in your bags under your desks or at the front of the room. Hands must be on top of your desks.*

v. *You are expected to simplify any expressions that can be simplified as whole numbers, e.g., $36/15 = 12/5$, but you may leave other expressions as is, e.g., $\sqrt{47}$ or $18/29$.*

vi. *You will be provided all of the critical values that you need.*

vii. *Show all your work to justify your answers. Answers without adequate justification will not receive credit.*

viii. *Print your name clearly on each answer sheet.*

ix. *Off-site students may contact the instructor as directed by your syllabus.*

Q	Score
1	
2	
3	
4	
Total	

I understand these instructions and have not relied on any help for this exam beyond what is allowed, nor provided any help to anyone else beyond what is allowed.

_____ _____

Signature *Date*

_____

Printed Name

Distribution	Sample Space	$\varrho(x)$	μ	σ^2
Poisson — Pois(λ)	$\{0, 1, 2, \ldots\}$	$\dfrac{\lambda^x \exp(-\lambda)}{x!}$	λ	λ

Upper quantiles: q_α satisfies $\mathbb{P}(X > q_\alpha) = \alpha$

X	α	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
Norm(0, 1)	z_α	-2.58	-2.33	-1.96	-1.64	-1.28	1.28	1.64	1.96	2.33	2.58
t_ν	$t_{24,\alpha}$	-2.80	-2.49	-2.06	-1.71	-1.32	1.32	1.71	2.06	2.49	2.80
	$t_{25,\alpha}$	-2.79	-2.49	-2.06	-1.71	-1.32	1.32	1.71	2.06	2.49	2.79
	$t_{26,\alpha}$	-2.78	-2.48	-2.06	-1.71	-1.31	1.31	1.71	2.06	2.48	2.78
F_{d_1,d_2}	$F_{1,47;\alpha}$	3.97e-05	0.000159	0.000993	0.00397	0.0160	2.82	4.05	5.36	7.21	8.68
	$F_{1,49;\alpha}$	3.97e-05	0.000159	0.000992	0.00397	0.0160	2.81	4.04	5.35	7.18	8.64
	$F_{2,47;\alpha}$	0.00501	0.0101	0.0253	0.0513	0.106	2.42	3.20	3.99	5.09	5.94
	$F_{2,49;\alpha}$	0.00501	0.0101	0.0253	0.0513	0.106	2.41	3.19	3.98	5.07	5.91

1. (25 points) Let $X_1, \ldots X_{25} \overset{\text{IID}}{\sim} \text{Norm}(\mu, \sigma^2)$. Furthermore, let

$$\overline{X} = \frac{1}{25}\sum_{i=1}^{25} X_i, \qquad \overline{x} = 12 = \text{sample mean from observed data,}$$

$$S^2 = \frac{1}{24}\sum_{i=1}^{25}(X_i - \overline{X})^2, \qquad s^2 = 100 = \text{sample variance from observed data.}$$

Fill in the blanks with the correct answers:

a. (5 points) $\overline{X}$ is an unbiased estimator of μ, and S^2 is an unbiased estimator of σ^2.

b. (5 points) A two-sided 95% confidence interval for the population mean based on the observed data is $\overline{x} \pm t_{24,0.025}s/\sqrt{25} = 12 \pm 2.06\sqrt{100/25} = 12 \pm 4.12 = [7.88, 16.12]$.

c. (5 points) The probability that a new IID observation, X_{26}, lies inside the interval in part b. is less than (less than, equal to, or greater than) 95%.

d. (5 points) *Before* collecting the data, it is claimed that population mean is 8.4 and you suspect that it is *larger* than 8.4. The hypothesis test that you would perform has null hypothesis $H_0 : \mu = 8.4$ and alternative hypothesis $H_A : \mu > 8.4$.

e. (5 points) Based on the observed data you would reject (reject or not reject) the null hypothesis at the 95% significance level.

Answer: Under the null hypothesis, statistic $T = (\overline{X} - 8.4)/(S/\sqrt{25}) \sim t_{24}$*, and the observed value is* $t = (\overline{x} - 8.4)/(s/\sqrt{25}) = 3.6/(10/5) = 1.8$*, which is greater than the critical value,* $t_{24,0.05} = 1.71$*, so we reject the null hypothesis.*

2. (24 points) The number of crashes of a server in a fixed time period is modeled by the Poisson distribution (see the table above). For IID data, $X_1, \ldots, X_n \overset{\text{IID}}{\sim} \text{Pois}(\lambda)$,

a. (7 points) Derive the likelihood function for λ.

Answer: The likelihood and log-likelihood functions are

$$L(\lambda \mid \boldsymbol{X}) = \prod_{i=1}^{n} \frac{\lambda^{X_i}\exp(-\lambda)}{X_i!} = \lambda^{X_1+\cdots+X_n}\exp(-n\lambda)\prod_{i=1}^{n}\frac{1}{X_i!}$$

$$\ell(\lambda \mid \boldsymbol{X}) = \log\big(L(\lambda \mid \boldsymbol{X})\big) = T\log(\lambda) - n\lambda - \sum_{i=1}^{n}\log(X_i!),$$

where $T = X_1 + \cdots + X_n$*.*

b. (7 points) Derive the maximum likelihood estimator for λ.

Answer: Setting the derivative of the log-likelihood function with respect to λ equal to zero gives

$$\frac{\partial \ell(\lambda \mid \boldsymbol{X})}{\partial \lambda} = \frac{T}{\lambda} - n$$
$$\widehat{\lambda}_{\text{MLE}} = \frac{T}{n} = \overline{X}, \text{ the sample mean}$$

c. (7 points) Show that $T = X_1 + \cdots + X_n$ is a *sufficient statistic* for λ.

Answer: Since the log likelihood may be written as

$$\ell(\lambda \mid \boldsymbol{X}) = g(T, \lambda) + h(\boldsymbol{X})$$

where $g(T, \lambda) = T \log(\lambda) - n\lambda$ and $h(\boldsymbol{X}) = -\sum_{i=1}^{n} \log(X_i!)$, it follows that T is a sufficient statistic.

d. (3 points) Derive the maximum likelihood estimator for $\text{var}(X_1)$.

Answer: Since $\text{var}(X_1) = \lambda$, $\widehat{\text{var}}(X_1) = \overline{X}$ as well.

3. (28 points) After fitting a regression model, $Y = \beta_1 + \beta_2 x + \beta_3 x^2 + \varepsilon$, to $n = 50$ data, computing the fitted response, $\widehat{y}_1, \ldots, \widehat{y}_{50}$, and computing the mean response, $\overline{y}$, you construct an ANOVA table using this data, but some values are missing.

Source	degrees of freedom	Sum of squares	Mean square	F
Regression	2	$\sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2 = 30$	$\frac{30}{2} = 15$	$\frac{\text{MSR}}{\text{MSE}} = \frac{15}{2} = 7.5$
Error	47	$\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 = 94$	$\frac{94}{47} = 2$	
Total	49	$\sum_{i=1}^{n}(y_i - \overline{y})^2 = 124$		

a. (15 points) Fill in the missing values in the ANOVA table.

b. (8 points) What does this ANOVA table tell you about your regression model? Be clear.

Answer: Because $F = 7.5 > F_{2,47;0.005} = 5.94$, we reject the null hypothesis that $\beta_2 = \beta_3 = 0$, so at least one predictor contributes significantly to the model.

c. (5 points) What is *one* thing that this ANOVA table cannot tell you about your regression model?

Answer:

- *Whether there are other variables that should be included*
- *Whether both x and x^2 should be in the model*
- *Whether the residuals vary with x or $\widehat{y}$*
- *Whether the residuals are normal*

4. (23 points) A class is tested on how many free throws they make out of 10 attempts before and after they receive training on shooting free throws. Let X be the number of baskets that an individual makes before training, Y be the number of baskets that an individual makes after training, and $D = Y - X$. We want to test whether *training helps.*

Here is some observed data.

Before, x_i	5	3	4	2	0	2	6	3	4	3	7	2
After, y_i	8	6	5	4	4	5	7	7	5	6	6	5

a. (7 points) Describe a possible *parametric* hypothesis test and explain why it would be unreliable.

Answer: We have a t-test for $H_0 : \mathbb{E}(D) = 0$ and $H_A : \mathbb{E}(D) > 0$, but this assumes that the D_i are normal, which does not fit this case.

b. (16 points) Perform an appropriate *nonparametric* test, stating the hypotheses, computing the p-value, and deciding whether the null hypothesis should be rejected at the 5% significance level.

Answer: We perform a sign test: $H_0 : \mathbb{P}(D > 0) = 1/2$ versus $H_A : \mathbb{P}(D > 0) > 1/2$. For this data we have

Before, x_i	5	3	4	2	0	2	6	3	4	3	7	2
After, y_i	8	6	5	4	4	5	7	7	5	6	6	5
Difference, $y_i - x_i$	3	3	1	2	4	3	1	4	1	3	−1	3
Rank	8	8	2.5	5	11.5	8	2.5	11.5	2.5	8	2.5	8

So, $s = \#\{y_i - x_i > 0\} = 11$. Let $S = \#\{D_i > 0\}$. Under the null hypothesis, the p-value is

$$\begin{aligned} p &= \mathbb{P}(S \geq s) = \mathbb{P}(S \geq 11) \\ &= \left[\binom{12}{11} + \binom{12}{12}\right]\left(\frac{1}{2}\right)^{12} \\ &= \left[\frac{12}{1} + 1\right]\frac{1}{4096} \\ &= \frac{13}{4096} \approx 0.00317 \end{aligned}$$

so we reject the null hypothesis in favor of the conclusion that training helps.

Another approach is the Wilcoxon signed-rank test with the hypotheses

$$H_0: \text{ the distribution of } D = Y - X \text{ is centered at } 0,$$
$$H_A: \text{ the distribution of } D = Y - X \text{ is centered above } 0.$$

The ranks are given above with tied ranks averaged.
Thus the sum of the negative signed ranks $W^- = 2.5$. *The total rank sum is* $1+2+\cdots+12 = 78$, *so the sum of the positive signed ranks is* $W^- = 78 - 2.5 = 75.5$.
For the one-sided alternative, large values of W^+, *equivalently small values of* W^-, *give evidence that training helps. Under* H_0, *the signs of the nonzero differences are equally likely to be positive or negative. The observed value* $W^- = 2.5$ *occurs when exactly one of the four observations with rank* 2.5 *is negative and all other differences are positive. The only more extreme case is* $W^- = 0$, *when all differences are positive. Therefore the p-value is*

$$p = \frac{1+4}{2^{12}} = \frac{5}{4096} \approx 0.00122.$$

Since $p < 0.05$, *we reject* H_0. *The data provide strong evidence that the training helps.*

Final Exam Scores
Number of Students: 23, Minimum: 22, Maximum: 84, Mean: 62, Median: 70
Standard Deviation: 17.9, Quartiles (Q1, Q3): (50, 75)

```
8|0234
7|00244559
6|1
5|02446
4|89
3|12
2|2
```