MATH 476 Statistics

Fred J. Hickernell | Test 1 | Thursday, February 19, 2026

Instructions:

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

ii. *The time allowed is 75 minutes.*

iii. *This test is closed book, but you may use 4 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.*

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
Geometric — Geom(p)	$\{1, 2, \ldots\}$	$p(1-p)^{x-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Uniform — Unif$[a, b]$	$[a, b]$	$\dfrac{1}{b-a}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$

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

1. (20 points) Given IID data, $X_1, \ldots, X_n$, what is the maximum likelihood estimator (MLE) for the parameter p in the geometric distribution, $\mathrm{Geom}(p)$ (see above)?

Answer: The likelihood and the log-likelihood functions in terms of the data, $\boldsymbol{X} = (X_1, \ldots, X_n)$ *are*

$$\begin{aligned} L(p \mid \boldsymbol{X}) &= \prod_{i=1}^{n} p(1-p)^{X_i - 1} = p^n (1-p)^{X_1 + \cdots + X_n - n} \\ \ell(p \mid \boldsymbol{X}) &= n \log(p) + [X_1 + \cdots + X_n - n] \log(1-p) \\ \frac{1}{n}\ell(p \mid \boldsymbol{X}) &= \log(p) + (\overline{X}_n - 1)\log(1-p), \qquad \overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \end{aligned}$$

Taking the derivative with respect to p *gives*

$$\begin{aligned} 0 &= \frac{1}{n}\frac{\partial \ell(p \mid \boldsymbol{X})}{\partial p} = \frac{1}{p} - \frac{\overline{X}_n - 1}{1-p} \\ \implies 0 &= 1 - p - p(\overline{X}_n - 1) = 1 - p\overline{X}_n \\ \implies p_{MLE} &= \frac{1}{\overline{X}_n} \end{aligned}$$

Many did well on this, but some were confused in computing the likelihood.

2. (44 points) A machine gives readings that are randomly distbributed $\mathrm{Unif}[\theta, \theta + 2]$ (see above) with unknown parameter θ.

   a. (8 points) If $\overline{X}_n$ is the sample mean of n IID samples, is $\overline{X}_n$ an unbiased estimator of θ? If not, modify it to make it one.

   Answer: For the uniform distribution, we know that $\mu = \mathbb{E}(X) = \theta + 1$ *and* $\sigma^2 = 4/12 = 1/3$. *The sample mean,* $\overline{X}_n$ *is an unbiased estimator of the population mean, which is* $\mu = \theta + 1$, *so* $\overline{X}_n$ *is a biased estimator of* θ, *but* $\overline{X}_n - 1$ *Is an unbiased estimator of* θ

   b. (20 points) Construct a large n approximate 95% confidence interval for θ in terms of $\overline{X}_n$.

   Answer: For large sample size, n, *we know that* $\overline{X}_n \dot\sim \mathrm{N}(\theta + 1, 1/(3n))$, *so*

$$\begin{aligned} 95\% &\approx \mathbb{P}\left[\theta + 1 - \frac{1.96}{\sqrt{3n}} \le \overline{X}_n \le \theta + 1 + \frac{1.96}{\sqrt{3n}}\right] \\ &= \mathbb{P}\left[\overline{X}_n - 1 - \frac{1.96}{\sqrt{3n}} \le \theta \le \overline{X}_n - 1 + \frac{1.96}{\sqrt{3n}}\right] \end{aligned}$$

   so the confidence interval is

$$\left[\overline{X}_n - 1 - \frac{1.96}{\sqrt{3n}} \le \theta \le \overline{X}_n - 1 + \frac{1.96}{\sqrt{3n}}\right]$$

c. (8 points) If $\overline{x}_{75} = 35$ is observed, then what is the half-width of this confidence interval?

Answer: The half width is $1.96/\sqrt{3 \times 75} = 1.96/\sqrt{225} = 1.96/15$

d. (8 points) How large would n need to be to make the confidence interval have a half-width of no greater than 0.1?

Answer: Setting half width to 0.1 *means that*

$$0.1 = \frac{1.96}{\sqrt{3n}} \iff n = \left(\frac{1.96}{0.1\sqrt{3}}\right)^2 = \frac{(19.6)^2}{3} \approx 128.1$$

So $n = 129$ *works.*

Some confused θ, μ and $\overline{X}_n$. Some constructed confidence interval for μ and not θ. Some confused the variance of the Bernoulli distribution with the variance of the Uniform distribution. Some did not simplify as far as they could without a calculator.

3. (24 points) Let

- $X_1, \ldots, X_n$ be IID miles driven until replacement for cars with Michelin tires,
- $Y_1, \ldots, Y_n$ be IID miles driven until replacement for an independent group of cars with Goodyear tires, and
- $D_i = X_i - Y_i$, for $i = 1, \ldots, n$

Let $\overline{X}$, $\overline{Y}$, and $\overline{D}$ and S_X^2, S_Y^2, and S_D^2, denote the respective sample means and variances of this random data.

Assuming large n, construct a 99% confidence interval in terms of this data for the correct population parameter that measures the difference between Michelin and Goodyear tires. Specify precisely the population parameter that your confidence interval is describing.

Answer: These two populations are independent *since the cars are not the same. This is* not *paired observations. Therefore we construct a confidence interval for the difference of two mean performances of the two brands of tires,* $\mu_X - \mu_Y$*. The* 99% *confidence interval uses the critical value* $z_{0.005} = 2.58$*, and we have*

$$\overline{X} - \overline{Y} \pm 2.58\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{n}}$$

It is important to specify what you are constructing a confidence interval for.

4. (12 points) You are an aide to the mayor. The mayor's draft speech states:

> We directed our city managers to ask their staff by a show of hands, "Do you approve of the mayor's leadership?" Out of 250 responses, 92.47% answered, "Yes." We conclude that 92,470 out of our city's 100,000 residents approve of my leadership.

Explain to the mayor at least three things wrong with this statistical inference.

Answer:

- *City workers are not an IID sample of the whole population?*
- *Non-anonymous responses are likely to be biased or inaccurate.*
- *You cannot have sample mean to four significant digits with a sample size of only* 250.
- *It is doubtful that the city's population is known to that accuracy, and*
- *One cannot know the opinions of babies, so one should speak about the* adult *population.*
- *The sample mean will usually not equal the population mean.*
- *A confidence interval for the mean will have a half-width of approximately* $1/\sqrt{250}$*, which is several percent.*

The sample size is large enough. It is just not IID.

Test 1 Scores (Regular and Make-up)

Number of Students: 24, Minimum: 18, Maximum: 100, Mean: 63.8, Median: 67.5
Standard Deviation: 22.6, Quartiles (Q1, Q3): (47.5, 84.5)

```
10|0
 9|02
 8|03666
 7|0246
 6|5
 5|12367
 4|34
 3|28
 2|7
 1|8
```