M378K Introduction to Mathematical Statistics
Fall 2025
University of Texas at Austin
Practice for the Final Exam
Instructor: Milica Čudina

Notes: This is a closed book and closed notes exam. The maximal score on the real exam will be 100 points.
There are many ways in which any single problem can be solved. The solutions herein are just one possible way to tackle the given problems.
Time: 120 minutes

All written work handed in by the student is considered to be
their own work, prepared without unauthorized assistance.

The University Code of Conduct

"The core values of The University of Texas at Austin are learning, discovery, freedom, leadership, individual opportunity, and responsibility. Each member of the university is expected to uphold these values through integrity, honesty, trust, fairness, and respect toward peers and community. As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity."

"I agree that I have complied with the UT Honor Code during my completion of this exam."

Signature:

4.1. **Formulas.** If Y has the binomial distribution with parameters n and p, then $p_Y(k) = \mathbb{P}[Y = k] = \binom{n}{k} p^k (1-p)^{n-k}$, for $k = 0, \dots, n$, $\mathbb{E}[Y] = np$, $\mathrm{Var}[Y] = np(1-p)$. The binomial coefficients are defined as follows for integers $0 \le k \le n$, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. The moment generating function of Y is given by $m_Y(t) = (pe^t + q)^n$.

If Y has a geometric distribution with parameter p, then $p_Y(k) = p(1-p)^k$ for $k = 0, 1, \dots$, $\mathbb{E}[Y] = \frac{1-p}{p}$, $\mathrm{Var}[Y] = \frac{1-p}{p^2}$. Its mgf is $m_Y(t) = \frac{p}{1-qe^t}$ for t such that $qe^t < 1$.

If Y has a Poisson distribution with parameter λ, then $p_Y(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ for $k = 0, 1, \dots$, $\mathbb{E}[Y] = \mathrm{Var}[Y] = \lambda$. Its mgf is $m_Y(t) = e^{\lambda(e^t - 1)}$.

If Y has a uniform distribution on $[l, r]$, its density is

$$f_Y(y) = \frac{1}{r-l} \mathbf{1}_{(l,r)}(y),$$

its mean is $\frac{l+r}{2}$, and its variance is $\frac{(r-l)^2}{12}$. Let $U \sim U(0,1)$. The mgf of U is $m_U(t) = \frac{1}{t}(e^t - 1)$.

If Y has the standard normal distribution, then its mean is zero, its variance is one, and its density equals

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad x \in \mathbb{R}.$$

Its mgf is $m_Y(t) = e^{\frac{t^2}{2}}$.

If Y has the exponential distribution with parameter τ, then its cumulative distribution function is $F_Y(y) = 1 - e^{-\frac{y}{\tau}}$ for $y \ge 0$, its probability density function is $f_Y(y) = \frac{1}{\tau} e^{-y/\tau}$ for $y \ge 0$. Also, $\mathbb{E}[Y] = SD[Y] = \tau$. Its mgf is $m_Y(t) = \frac{1}{1-\tau t}$.

The mgf of $Y \sim \Gamma(k, \tau)$ is

$$m_Y(t) = \frac{1}{(1-\tau t)^k} \text{ for } t < 1/\tau.$$

Its expectation is $k\tau$ and its variance is $k\tau^2$. The χ^2-distribution with n degrees of freedom is the special case $\Gamma\left(\frac{n}{2}, 2\right)$

4.2. **DEFINITIONS. You should know all the definitions from previous exams.**

4.3. **TRUE/FALSE QUESTIONS.**

Problem 4.1. (2 points) The smaller the $p-$value, the stronger the evidence that the null hypothesis should be rejected. *True or false? Why?*

Solution: TRUE
Since the $p-$value is the probability that the observed value of the test statistic or something more extreme occurred under the null hypothesis.

Problem 4.2. (2 points) The $p-$value denotes the probability that the null hypothesis is rejected given that the null hypothesis is true. *True or false? Why?*

Solution: FALSE
This is actually the **significance level**.

4.4. **Free-response problems.**

Please, explain carefully all your statements and assumptions. Numerical results or single-word answers without an explanation (even if they're correct) are worth 0 points.

Problem 4.3. Consider an election between exactly two candidates: A and B. Let p denote the population proportion of voters in favor of candidate A. We wish to test

$$H_0 : p = \frac{1}{2} \quad \text{vs.} \quad H_a : p < \frac{1}{2}.$$

A sample of 16 voters will be polled. Based on the above and a specific significance level α, Bertie Wooster obtained the rejection region of the form $RR = [0, 2]$. Of course, he then forgot α. Can you help Bertie by calculating the lowest α that could have corresponded to the above rejection region? You should leave your response in the form of a completely reduced fraction.

Solution: Let T denote the sample count of people who said they would vote for candidate A. Then, the sampling distribution of T is, under the null hypothesis,

$$T \sim B\left(n = 16, p = \tfrac{1}{2}\right)$$

Our α must satisfy

$$\alpha = \mathbb{P}[T = 0] + \mathbb{P}[T = 1] + \mathbb{P}[T = 2] = \left(\frac{1}{2}\right)^{16} + 16\left(\tfrac{1}{2}\right)^{16} + \binom{16}{2}\left(\frac{1}{2}\right)^{16} = \frac{1 + 16 + 120}{2^{16}} = \frac{137}{2^{16}}(= 0.002090454$$

Problem 4.4. *Source: "Mathematical Statistics with Applications" by Wackerly, Mendenhall and Scheaffer.*
A vice president in charge of sales for a large corporation claims that salespeople are averaging no more than 15 sales contacts per week. (He would like to increase this figure.) We are interested in the research hypothesis that the vice president's claim is incorrect. As a check on his claim, $n = 36$ salespeople are selected at random, and the number of contacts made by each is recorded for a single randomly selected week. The mean and variance of the 36 measurements were 17 and 9, respectively. Does the evidence contradict the vice president's claim? Use a test with significance level $\alpha = 0.05$. Make sure that you state all the stages of a hypothesis test and substantiate all your claims.

Solution: Let μ denote the population mean, i.e., the mean number of weekly sales contacts in the sales force. We are testing

$$H_0 : \mu = 15 \quad \text{against} \quad H_0 : \mu > 15.$$

With $\bar{Y}$ denoting the sample mean and S denoting the sample variance, our test statistic is, under the null hypothesis,

$$T = \frac{\bar{Y} - \mu_0}{\frac{S}{\sqrt{n}}}$$

Under the null, its sampling distribution is t with 35 degrees of freedom. The observed value of our test statistic is

$$t = \frac{17 - 15}{\frac{3}{\sqrt{36}}} = 4.$$

With 35 degrees of freedom, the $t-$distribution is close to being standard normal. So, the $p-$value, i.e., the probability $\mathbb{P}[T > t]$ is much much smaller than the significance level $\alpha = 0.05$. We reject the null hypothesis. The data point to the manager being wrong.

Problem 4.5. A special loaded coin is claimed to be calibrated so that it lands on *heads* with the chance $\frac{9}{13}$ and *tails* with the chance $\frac{4}{13}$. You suspect that the coin is incorrectly calibrated and you conduct an appropriate hypothesis test. Your experiment consists of 64 independent tosses of this coin. You obtain 48 *heads*. What is your verdict at the 0.02 significance level? Make sure that you state all the stages of a hypothesis test and substantiate all your claims.

Solution: Let p denote the probability that the coin lands on *heads*. We should test

$$H_0 : p = \tfrac{9}{13} \quad \text{against} \quad H_a : p \neq \tfrac{9}{13}\,.$$

Since there are 64 trials, we can use the $z-$procedure. In this case, the appropriate test statistic is

$$T = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

which is approximately standard normal under the null hypothesis.

The observed sample proportion is $\hat{p} = \frac{48}{64} = \frac{3}{4}$. So, the observed value of our test statistic is

$$t = \frac{\frac{3}{4} - \frac{9}{13}}{\sqrt{\frac{\frac{9}{13}\left(\frac{4}{13}\right)}{64}}} = \frac{\frac{39-36}{4\cdot 13}}{\frac{3\cdot 2}{13\cdot 8}} = 1.$$

The $p-$value is

$$2\Phi\left(-1\right) > 0.02.$$

Evidently, we fail to reject the null hypothesis.