M378K Introduction to Mathematical Statistics
Fall 2025
University of Texas at Austin
Practice for In-Term Exam II
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: 50 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:

2.1. **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 $0 \leq k \leq n$, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. 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 $y \geq 0$, its probability density function is $f_Y(y) = \frac{1}{\tau} e^{-y/\tau}$ for $y \geq 0$. 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 $\Gamma\left(\frac{n}{2}, 2\right)$

2.2. DEFINITIONS.

Problem 2.1. (10 points) Write down the definition of the **moment generating function** of a random variable Y.

Solution:

$$m_Y(t) = \mathbb{E}[e^{tY}]$$

for all $t \in \mathbb{R}$ such that the above expectation exists. In fact, we say that the moment-generating function **exists** if there exists a positive number b such that $m_Y(t)$ is finite for all t such that $|t| \leq b$.

2.3. TRUE/FALSE QUESTIONS.

Problem 2.2. (5 points) The random vector (X, Y) is jointly continuous with the joint probability density function given by

$$f_{(X,Y)}(x, y) = \begin{cases} (1/8)xe^{-(x+y)/2}, & x > 0, y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then, the random variables X and Y are independent. *True or false? Why?*

Solution: TRUE
The joint p.d.f. can be rewritten as

$$f_{(X,Y)}(x, y) = \frac{1}{4}xe^{-x/2} \times \frac{1}{2}e^{-y/2} = f_X(x)f_Y(y)\,.$$

So, the criterion for independence of jointly continuous random variables is satisfied. We conclude that X and Y are independent.

2.4. 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 2.3. (15 points) Let Y_1 and Y_2 be independent exponential random variables with parameters τ_1 and τ_2.

(1) (5 points) What is the joint density of (Y_1, Y_2)?
(2) (10 points) Compute $\mathbb{P}[Y_1 \geq Y_2]$.

Solution:

(1) Since Y_1 and Y_2 are independent, the joint density of the pair (Y_1, Y_2) is the product of the marginal densities $f_{Y_1}(y) = \frac{1}{\tau_1}e^{-y/\tau_1}\mathbf{1}_{\{y>0\}}$ and $f_{Y_2}(y) = \frac{1}{\tau_2}e^{-y/\tau_2}\mathbf{1}_{y>0}$ evaluated in the variables y_1 and y_2, respectively, i.e.,

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{\tau_1}e^{-y_1/\tau_1}\mathbf{1}_{\{y_1>0\}} \times \frac{1}{\tau_2}e^{-y_2/\tau_2}\mathbf{1}_{\{y_1>0\}} = \tfrac{1}{\tau_1\tau_2}e^{-y_1/\tau_1 - y_2/\tau_2}\mathbf{1}_{\{y_1,y_2\geq 0\}}.$$

(2) To compute $\mathbb{P}[Y_1 \geq Y_2]$, we need to integrate f_{Y_1,Y_2} over the region $A = \{(y_1, y_2) \in \mathbb{R}^2 : y_1 \geq y_2\}$, which we immediately turn into an indicator $\mathbf{1}_{\{y_1 \geq y_2\}}$:

$$\begin{aligned}\mathbb{P}[Y_1 \geq Y_2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{1}_{\{y_1\geq y_2\}} \frac{1}{\tau_1\tau_2} e^{-y_1/\tau_1 - y_2/\tau_2} \mathbf{1}_{\{y_1,y_2\geq 0\}}\, dy_2\, dy_1 \\ &= \int_0^{\infty}\int_0^{y_1} \frac{1}{\tau_1\tau_2} e^{-y_1/\tau_1 - y_2/\tau_2}\, dy_2\, dy_1 \\ &= \int_0^{\infty} \frac{1}{\tau_1} e^{-y_1/\tau_1}(1 - e^{-y_1/\tau_2})\, dy_1 = \tfrac{\tau_1}{\tau_1+\tau_2}.\end{aligned}$$

Problem 2.4. (10 points) Let $Y \sim U(l, r)$ What is the moment generating function of Y?

Solution: See **Example 6.1.5** from the lecture notes.

Problem 2.5. (20 points) Luka owns a mechanical pencil. The lifetime of every piece of lead has mean 4 days and standard deviation of 1 day. The lifetimes of different lead pieces are independent. After one lead piece is exhausted, Luka immediately uses the next one. What is the smallest number of packets (each containing 10 pieces of lead) that Luka should buy in order to have enough lead for the next 360 days with probability at least 0.9987?

Hint:

$$\sqrt{(8(360)+9)^2 - 4(16)(360^2)} \approx 228$$

Solution: Let n denote the total number of pieces of lead Luka should buy. Let $Y_i, i = 1, \ldots, n$ be the random variables denoting the lifetimes of individual lead pieces. Then, their total lifetime is

$$S = Y_1 + \cdots + Y_n$$

We suspect that we can use the Central Limit Theorem (CLT) here since the total time given is 360 while an individual piece of lead survives, on average, for 4 days. Let's see what we get with the CLT. We would have that S is approximately normal with mean $4n$ and variance n.

Our condition from the problem can be expressed as

$$\mathbb{P}[S > 360] \geq 0.9987 \quad \Leftrightarrow \quad \mathbb{P}\left[\frac{S-4n}{\sqrt{n}} > \frac{360-4n}{\sqrt{n}}\right] \geq 0.9987.$$

With the CLT, we get

$$\mathbb{P}\left[Z > \frac{360-4n}{\sqrt{n}}\right] \geq 0.9987$$

where $Z \sim N(0, 1)$. This is equivalent to saying

$$\mathbb{P}\left[Z \leq \frac{360-4n}{\sqrt{n}}\right] \leq 0.0013.$$

If we consult the standard normal tables, we find that the equivalent condition for the borderline n is

$$\frac{360-4n}{\sqrt{n}} = -3 \quad \Leftrightarrow \quad (360-4n) = -3\sqrt{n}.$$

This leads us to the quadratic

$$(360-4n)^2 = 9n, \quad \text{i.e.,} \quad 360^2 - 8(360)n + 16n^2 = 9n, \quad \text{i.e.,} \quad 16n^2 - (8(360)+9)n + 360^2 = 0.$$

We get two solutions

$$n_{\pm} = \frac{(8(360)+9) \pm \sqrt{(8(360)+9)^2 - 4(16)(360^2)}}{2(16)} = \frac{8(360)+9 \pm 228}{32}.$$

We are going to keep the **bigger** of the two solutions (the other one is symmetric on the bell curve with respect to the mean). We get

$$n = \frac{8(360)+9+228}{32} = \frac{3117}{32} > 90.$$

So, Luka needs to buy at least 10 boxes of pencil leads.

Problem 2.6. (20 points) In Croatia, if you go to the chocolate-factory store, you can buy broken off chunks of rice-puff chocolate. From past experience, we know that the weight of the individual chunks has mean of 40 grams and standard deviation of 5 grams. Assume that the weights of individual pieces of chocolate are independent.

You buy 400 chocolate chunks. What is the probability that the total weight exceeds 16128 grams?

Solution: Let $n = 400$ denote the total number of chocolate chunks. Let $Y_i, i = 1, \dots, n$ be the random variables which stand for the weights of individual chunks. Then, their total weight can be expressed as

$$S = Y_1 + \cdots + Y_n$$

We can use the Central Limit Theorem (CLT) here since $n = 400$. We have that S is approximately normal with mean $40(400) = 16000$ and standard deviation $5\sqrt{400} = 100$.

The probability we are asked to calculate is

$$\mathbb{P}[S > 16128] = \mathbb{P}\left[\frac{S - 16000}{100} > \frac{16128 - 16000}{100}\right] \approx \mathbb{P}\left[Z > 1.28\right]$$

where $Z \sim N(0,1)$. We get

$$\mathbb{P}[S > 16128] \approx 1 - \Phi(1.28).$$

If we consult the standard normal tables, we get our answer as $1 - 0.8997 = 0.1003$.

2.5. MULTIPLE CHOICE QUESTIONS.

Problem 2.7. (5 points) The pdf of $W = 1/Y^2$, where $Y \sim E(\tau)$, is ...

(a) $\frac{2}{\tau} y^{-3/2} e^{-y/\tau} \mathbf{1}_{\{y>0\}}$
(b) $\frac{1}{2\tau}(-y^{-3/2}) e^{-\sqrt{y}/\tau} \mathbf{1}_{\{y>0\}}$
(c) $\frac{1}{\tau} e^{-1/(y^2\tau)} \mathbf{1}_{\{y>0\}}$
(d) $\frac{1}{2\tau y^{3/2}} e^{-1/(\tau\sqrt{y})} \mathbf{1}_{\{y>0\}}$
(e) none of the above

Solution: The correct answer is **(d)**.
Since $g(y)$ is continuously differentiable and decreasing on $(0, \infty)$ and $f_Y(y) = 0$ for $y \leq 0$, we can use the h-method with $h(w) = 1/\sqrt{w}$:

$$f_W(w) = f_Y(h(w))\left|h'(w)\right| = \tfrac{1}{\tau}e^{-1/(\tau\sqrt{w})}\tfrac{1}{2}w^{-3/2}\mathbf{1}_{\{w>0\}}.$$

Problem 2.8. (5 points) Let $Y_1, Y_2, \ldots, Y_n$ be independent, standard normal random variables. What is the distribution of the random variable Y defined as

$$Y = Y_1^2 + Y_2^2 + \cdots + Y_n^2\,?$$

(a) $N(0, \sqrt{n})$
(b) $\chi^2(n)$
(c) $\chi^2(n-1)$
(d) $N(0, n^2)$
(e) **None of the above.**

(*Note:* In our notation $N(\mu, \sigma)$ means normal with mean μ and *standard deviation* σ.)

Solution: The correct answer is **(b)**.

Problem 2.9. (5 points) *Source: Sample P exam, Problem #250.*
A delivery service owns two cars that consume 15 and 20 miles per gallon, respectively. Fuel costs \$3 per gallon. On any given business day, each car travels a number of miles that is independent of the other and is normally distributed with mean 25 miles and standard deviation 3 miles. Calculate the probability that on any given business day, the total fuel cost to the delivery service will be less than 7.

(a) About 0.0013
(b) About 0.0073
(c) About 0.0099
(d) About 0.0138
(e) **None of the above.**

Solution: The correct answer is **(c)**.
Given the efficiencies of the two cars, we have that their daily fuel consumption can be modeled by

$$Y_1 \sim N\left(\mu_1 = \tfrac{25}{15} = \tfrac{5}{3}, \sigma_1 = \tfrac{3}{15} = \tfrac{1}{5}\right) \quad \text{and} \quad Y_2 \sim N\left(\mu_2 = \tfrac{25}{20} = \tfrac{5}{4}, \sigma_2 = \tfrac{3}{20}\right).$$

Hence, the total fuel consumption will be

$$Y = Y_1 + Y_2 \sim N(\mu = \mu_1 + \mu_2 = \tfrac{35}{12}, \sigma = \sqrt{\sigma_1^2 + \sigma_2^2} = 0.25).$$

Taking into account the cost of fuel per gallon, the we need to find the probability

$$\mathbb{P}[3Y < 7] = \mathbb{P}\left[\frac{3Y - 3\mu}{3\sigma} < \frac{7 - 3\mu}{3\sigma}\right] = \mathbb{P}\left[Z \leq \frac{7 - 8.75}{0.75}\right] \approx \mathbb{P}[Z \leq -2.33]$$

where $Z \sim N(0.1)$. So, our answer is $\Phi(-2.33) = 0.0099$.

Problem 2.10. (5 points) *Source: Sample P exam, Problem #362.*
At a certain airport, $\frac{1}{6}$ of all scheduled flights are delayed. Assume that flight delays are mutually independent events. Use the normal approximation **with continuity correction** to calculate the probability that at least 40 of the next 180 flights are delayed.

(a) About 0.0110
(b) About 0.0143
(c) About 0.0182
(d) About 0.0234
(e) About 0.0287

Solution: The correct answer is **(e)**.
The normal approximation to the binomial is applicable. We have that the exact distribution of the number of delayed flights is

$$Y = B\left(n = 180, p = \tfrac{1}{6}\right)$$

To use the normal approximation, we need

$$\mu = np = 180\left(\frac{1}{6}\right) = 30, \quad \text{and} \quad \sigma = \sqrt{npq} = \sqrt{180\left(\tfrac{1}{6}\right)\left(\tfrac{5}{6}\right)} = 5$$

Hence, the approximate probability can be evaluated as

$$\mathbb{P}[Y \geq 40] \approx 1 - \Phi\left(\frac{39.5 - 30}{5}\right) = 1 - \Phi(1.90) = 0.0287.$$