Name:

UTeid:

M378K Introduction to Mathematical Statistics
Fall 2025
University of Texas at Austin
In-Term Exam II
Instructor: Milica Čudina

Notes: This is a closed book and closed notes exam. The maximal score on the exam is 100 points.
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 $0 \le k \le n$, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. 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 $m_Y(t) = \frac{1}{1-\tau t}$ on $|t| < \tau^{-1}$.

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) Let the random vector (Y_1, Y_2) have the following joint probability density function:

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} \frac{3}{\sqrt{2\pi}} \frac{y_2^{1/2}}{y_1^4} e^{-y_2/2}, & y_1 \geq 1,\ y_2 > 0, \\ \\ 0, & \text{otherwise.} \end{cases}$$

Then, the random variables Y_1 and Y_2 are independent. *True or false? Why?*

Solution: TRUE
We can use the factorization criterion with

$$g(y_1) = \frac{3}{\sqrt{2\pi}} y_1^{-4} \mathbf{1}_{[1,\infty)}(y_1) \tag{2.1}$$

and

$$g(y_2) = y_2^{1/2} e^{-y_2/2} \mathbf{1}_{[0,\infty)}(y_2) \tag{2.2}$$

Problem 2.3. (5 points) Let Y_1 be Poisson with mean 1 and let Y_2 be an independent Poisson with mean 3. Then, $Y = \frac{1}{2}(Y_1 + Y_2)$ is Poisson with mean 2. *True or false? Why?*

Solution: FALSE
The support is wrong (among other things).

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 2.4. (20 points) In a local candy store, you can buy irregular blobs of *Freaky Fudge*. The weight of each blob is a random variable with a mean of 50 grams and a standard deviation of 10 grams. Assume the weights of individual blobs are independent.

You buy 144 blobs of fudge for your friends and relations. What is the probability that the total weight exceeds 6900 grams?

Solution: Let $n = 144$ denote the total number of fudge blobs. Let $Y_i, i = 1, \dots, n$ be the random variables which stand for the weights of individual blobs. Then, their total weight can be expressed as

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

We can use the Central Limit Theorem (CLT) here since $n = 144$. We have that S is approximately normal with mean $50(144) = 7200$ and standard deviation $10\sqrt{144} = 120$.

The probability we are asked to calculate is

$$\mathbb{P}[S > 6900] = \mathbb{P}\left[\frac{S - 7200}{120} > \frac{6900 - 7200}{120}\right] \approx \mathbb{P}\left[Z > -2.5\right]$$

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

$$\mathbb{P}[S > 6900] \approx \Phi(2.5).$$

If we consult the standard normal tables, we get our answer as 0.9938.

Problem 2.5. (10 points) A dart player throws a dart at a dartboard - the board itself is always hit, but any region of the board is as likely to be hit as any other of the same area. We model the board as the unit disc $\{y_1^2 + y_2^2 \leq 1\}$, and the point where the board is hit by a pair of random variables (Y_1, Y_2). This means that (Y_1, Y_2) is uniformly distributed on the unit disc, i.e., the joint pdf is given by

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{\pi}\mathbf{1}_{\{y_1^2+y_2^2\leq 1\}} = \begin{cases} \frac{1}{\pi}, & y_1^2 + y_2^2 \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

(i) (5 points) Are the random variables Y_1 and Y_2 to be independent? Explain why or why not (do **not** do any calculations).

(ii) (5 points) Compute $\mathbb{P}[Y_1 \geq Y_2]$.

Solution:

(i) Without any knowledge about the value of Y_2, we cannot rule out any any value in the interval $[-1, 1]$ as a possible value of Y_1. We know, however, it will always be the case that $Y_1^2 \leq 1-Y_2^2$. Therefore, if the value of Y_2 is revealed to be, say, $3/5$, the value of $|Y_1|$ can no longer be larger than $\sqrt{1-(3/5)^2} = 4/5$. Therefore, we would expect Y_1 and Y_2 not to be independent.

(ii) $A = \{(y_1, y_2) \in \mathbb{R}^2 : y_1 \geq y_2\}$ is the region under (to the right) of the line $y_1 = y_2$. Since the distribution of (Y_1, Y_2) is uniform on the unit disc $D = \{(y_1, y_2) \in \mathbb{R}^2 : y_1^2 + y_2^2 \leq 1\}$, the required probability is given by the quotient of the areas of $D \cap A$ and D. The region $D \cap A$ is a half-disc. From this we conclude that the probability is given by

$$\frac{\text{area}(A\cap D)}{\text{area}(D)} = \frac{\pi/2}{\pi} = \frac{1}{2}.$$

Problem 2.6. (10 points) Let Y_1 be exponential with mean τ_1 and let Y_2 be exponential with mean τ_2. Assume that Y_1 and Y_2 are independent. What is the moment generating function of $Y = Y_1 + Y_2$? Do not forget to state the domain!

Solution: The moment generating functions of the given random variables are

$$m_{Y_1}(t) = \frac{1}{1 - \tau_1 t} \quad \text{and} \quad m_{Y_2}(t) = \frac{1}{1 - \tau_2 t}$$

with the domains consisting of all t such that $|t| < \frac{1}{\tau_1}$ and $|t| < \frac{1}{\tau_2}$. Then, the moment generating of $Y = Y_1 + Y_2$ is

$$m_Y(t) = m_{Y_1}(t) m_{Y_2}(t) = \left(\frac{1}{1 - \tau_1 t}\right)\left(\frac{1}{1 - \tau_2 t}\right) = \frac{1}{(1 - \tau_1 t)(1 - \tau_2 t)}$$

with the domain $|t| < \min\left(\frac{1}{\tau_1}, \frac{1}{\tau_2}\right)$.

Problem 2.7. (15 points) A spider starts off at a central position up a spout and wanders randomly up and down. Its position after a while - measured from the perspective of the spout - is modeled as normal with mean (naturally) zero and standard deviation 3. Let Y denote the final distance of the spider from its initial position. What is the density of Y expressed in terms of the standard normal probability density function φ? Simplify your expression as much as possible.

Solution: Let X denote the spider's position. We know that

$$X \sim N(0, \sigma = 2) \quad \text{and} \quad Y = |X|$$

Let's start with the cdf of Y. For all $y > 0$, we have

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \leq y] = \mathbb{P}[|X| \leq y] \\ &= \mathbb{P}[-y \leq X \leq y] = \mathbb{P}[X \leq y] - \mathbb{P}[X \leq -y] \\ &= \mathbb{P}\left[\frac{X}{3} \leq \frac{y}{3}\right] - \mathbb{P}\left[\frac{X}{3} \leq -\frac{y}{3}\right] = \mathbb{P}\left[Z \leq \frac{y}{3}\right] - \mathbb{P}\left[Z \leq -\frac{y}{3}\right] \end{aligned}$$

where $Z \sim N(0,1)$. Now, in terms of the standard normal cumulative distribution function, we have

$$F_Y(y) = \Phi\left(\frac{y}{3}\right) - \Phi\left(-\frac{y}{3}\right) = 2\Phi\left(\frac{y}{3}\right) - 1.$$

Finally, the pdf of Y is

$$f_Y(y) = \frac{2}{3}\varphi\left(\frac{y}{3}\right) \mathbf{1}_{[0,\infty)}(y).$$

2.5. **MULTIPLE CHOICE QUESTIONS.**

Problem 2.8. (5 points) Let Y be a uniform random variable on $[0,1]$, and let $W = Y^4$. The pdf of W is

(a) $\frac{1}{2\sqrt{|w|}}\mathbf{1}_{\{-1<w<1\}}$

(b) $\frac{1}{4}w^{-3/4}\mathbf{1}_{\{0<w\leq 1\}}$

(c) $\frac{1}{2\sqrt{w}}\mathbf{1}_{\{0<w<1\}}$

(d) $4w\mathbf{1}_{\{0<w<1\}}$

(e) **None of the above.**

Solution: The correct answer is **(b)**.
We use the h-method. $g(y) = y^4$ and $h(w) = w^{1/4}$. Therefore

$$f_W(w) = f_Y(h(w))h'(w) = \frac{1}{4}w^{-3/4}\mathbf{1}_{\{0<w\leq 1\}}.$$

Problem 2.9. (5 points) Let $Y_1, Y_2, \ldots, Y_n$ be independent, identically distributed normal random variables with mean μ and standard deviation σ. What is the distribution of the random variable Y defined as

$$Y = \left(\frac{Y_1 - \mu}{\sigma}\right)^2 + \left(\frac{Y_2 - \mu}{\sigma}\right)^2 + \cdots + \left(\frac{Y_n - \mu}{\sigma}\right)^2 ?$$

(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 2.10. (5 points) A math graduate student basically survives on espresso and chocolate. Their daily chocolate consumption is normally distributed with mean 16 oz and standard deviation 4 oz. Their daily espresso consumption is normally distributed with mean 20 oz and standard deviation 3 oz. Assume that espresso consumption and chocolate consumption are independent.

Let Φ denote the standard normal cumulative distribution function. What is the probability that chocolate consumption exceeds coffee consumption in a single day in terms of Φ?

(a) $\Phi(4/7)$
(b) $\Phi(4/5)$
(c) $1 - \Phi(4/5)$
(d) $1 - \Phi(4/7)$
(e) **None of the above.**

Solution: The correct answer is **(c)**.
Let Y_1 be the chocolate consumption and let Y_2 be the espresso consumption. We are given that Y_1 and Y_2 are independent. Also,

$$Y_1 \sim N(\mu_1 = 16, \sigma_1 = 4) \quad \text{and} \quad Y_2 \sim N(\mu_2 = 20, \sigma_2 = 3).$$

We need to calculate $\mathbb{P}[Y_1 > Y_2] = \mathbb{P}[Y_1 - Y_2 > 0]$. From the given information, we can conclude that

$$Y_1 - Y_2 \sim N(\mu = -4, \sigma = \sqrt{3^2 + 4^2} = 5).$$

So,

$$\mathbb{P}[Y_1 - Y_2 > 0] = \mathbb{P}\left[\frac{Y_1 - Y_2 - \mu}{\sigma} > \frac{0 - \mu}{\sigma}\right] = \mathbb{P}[Z > 0.8] = 1 - \Phi\left(\frac{4}{5}\right)$$

where $Z \sim N(0,1)$. By the way, from the standard normal tables, we get $1 - 0.7881 = 0.2119$.

Problem 2.11. (5 points) Let $Y_1, \dots, Y_{100}$ be independent random variables with the Bernoulli $B(p)$ distribution, with $p = 0.2$ The best approximation to $\bar{Y} = \frac{1}{n}(Y_1 + \dots + Y_n)$ (among the offered answers) is

(a) $N(0, 1)$
(b) $N(100, 20)$
(c) $N(0.2, 0.04)$
(d) $N(20, 4)$
(e) $N(20, 20)$

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

Solution: The correct answer is **(c)**.
The sum $W = Y_1 + \dots + Y_n$ is binomially distributed with mean $np = 20$ and variance $np(1-p) = 16$, i.e., standard deviation 4. It is well approximated by a normal $N(20, 4)$. Since $\bar{Y} = \frac{1}{n}W$, its best approximation will a normal with mean $\frac{1}{100}20 = 0.2$ and standard deviation $\sigma = \frac{1}{100}4 = 0.04$.

Problem 2.12. (5 points) Let (Y_1, Y_2) be a random vector with the following distribution table

	-1	1
1	$\frac{1}{6}$	$\frac{1}{2}$
2	$*$	$\circ$

If it is known that Y_1 and Y_2 are independent, the values $*$ and $\circ$ in the second row are

(a) $* = 1/6$, $\circ = 3/4$
(b) $* = 1/12$, $\circ = 1/4$
(c) $* = 1/6$, $\circ = 1/6$
(d) $* = 1/24$, $\circ = 7/24$
(e) **None of the above.**

Solution: The correct answer is **(b)**.
Let $p_1 = \mathbb{P}[Y_1 = -1]$ and $p_2 = \mathbb{P}[Y_2 = 1]$. The independence and the entries in the table yield the following equations for p_1 and p_2:

$$\tfrac{1}{6} + \tfrac{1}{2} = p_2, \tfrac{1}{6} = p_1 p_2.$$

It follows that $p_2 = 2/3$ and $p_1 = 1/4$. Therefore $* = p_1(1-p_2) = 1/12$ and $\circ = (1-p_1)(1-p_2) = 1/4$.
