Midterm Exam Solutions

Introduction to Statistics (통계의 이해) **2026 1st Semester**

Year(학년): ___	Section(교반): ___	Cadet Number(교번): _______	Name(성명): _________	Score: _______

- **All solutions must include a detailed step-by-step explanation.**
- **If an answer has five or more decimal places, round to the fourth decimal place.**
- **Use the reference materials provided on the last page of the exam.**

1. Read the following and fill in each blank with the appropriate term. (Use Reference Table 1.) [8 points]

(1) The probability distribution function of a discrete random variable X, $f(x) = P(X = x)$, is called the ( ) of X.

(2) When a Bernoulli trial with success probability p is repeated n times independently, the number of successes X follows the ( ) distribution.

(3) A ( ) is a function of a random sample that does not involve unknown parameters.

(4) If $Z_1, Z_2, \ldots, Z_n \sim N(0,1)$ independently, then $V := Z_1^2 + Z_2^2 + \cdots + Z_n^2$ follows a chi-squared distribution with ( ) degrees of freedom.

(5) A statistic $\hat{\theta}(X_1, \ldots, X_n)$ used to estimate a parameter θ is called an *estimator*, and the computed value $\hat{\theta}(x_1, \ldots, x_n)$ from observed data $(x_1, \ldots, x_n)$ is called the ( ).

(6) An estimator $\hat{\theta}$ of parameter θ is ( ) if $E(\hat{\theta}) = \theta$.

(7) If $X_1, \ldots, X_n$ is a random sample from a population with mean μ and variance σ^2, then the sample variance $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is both an unbiased and consistent estimator of ( ).

(8) If a point estimate follows a normal distribution, a $100(1-\alpha)\%$ ( ) for the population parameter is

$$\text{point estimate} \pm z_{\alpha/2} \times SE,$$

where $z_{\alpha/2}$ is the critical value and SE is the standard error.

Solution: (1) probability mass function / (2) binomial / (3) statistic / (4) n / (5) point estimate / (6) unbiased / (7) population variance σ^2 / (8) confidence interval

2. A report states that the probability of the enemy launching an attack this month is 0.01. A radar detection system is used to identify incoming signals. When an attack occurs, the system detects an alarm with probability 0.98. When no attack occurs, the system still detects an alarm with probability 0.05. Answer the following questions. [8 points]

(1) Find the probability that the radar detection system detects an alarm. (Define the relevant events before solving.)

Solution: Let A be the event that the enemy launches an attack, and B be the event that the radar system detects an alarm. Then

$$P(A) = 0.01, \quad P(A^c) = 0.99, \quad P(B \mid A) = 0.98, \quad P(B \mid A^c) = 0.05.$$

Since A and A^c form a partition of the sample space, by the law of total probability,

$$\begin{aligned} P(B) &= P(A)P(B \mid A) + P(A^c)P(B \mid A^c) \\ &= 0.01 \times 0.98 + 0.99 \times 0.05 = 0.0593. \end{aligned}$$

(2) Given that the radar detection system has detected an alarm, find the probability that the enemy actually launched an attack.

Solution: Since A and A^c form a partition of the sample space, by Bayes' theorem,

$$\begin{aligned} P(A \mid B) &= \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A^c)P(B \mid A^c)} \\ &= \frac{0.01 \times 0.98}{0.01 \times 0.98 + 0.99 \times 0.05} = \frac{0.0098}{0.0593} = 0.1653. \end{aligned}$$

Midterm Exam Solutions

Introduction to Statistics (통계의 이해) **2026 1st Semester**

Year(학년): ___	Section(교반): ___	Cadet Number(교번): _______	Name(성명): _________	Score: _______

3. A biased coin is tossed $n = 3$ times, where the probability of heads is $p = \dfrac{1}{3}$. Let X be the number of heads and Y be the number of tails. Answer the following questions. [12 points]

(1) Find the probability mass function $f_X(x)$ of X.

Solution: Since $X \sim B(3, \frac{1}{3})$,

$$f_X(x) = \begin{cases} \binom{3}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{3-x}, & x = 0, 1, 2, 3, \\ 0, & \text{otherwise.} \end{cases}$$

(2) The joint pmf of X and Y is given in the table below. Find the values ① and ②.

$x \setminus y$	**0**	**1**	**2**	**3**
0	0	0	0	①
1	0	0	12/27	0
2	0	6/27	0	0
3	1/27	②	0	0

Solution: ① $P(X = 0, Y = 3) = \left(\dfrac{2}{3}\right)^3 = \dfrac{8}{27}$ ② $P(X = 3, Y = 1) = 0$

(3) Determine whether X and Y are independent.(Hint: the probability mass function of Y is $f_Y(y) = \binom{3}{y}\left(\frac{2}{3}\right)^y\left(\frac{1}{3}\right)^{3-y}, \quad y = 0, 1, 2, 3.$)

Solution: Note that $f(1, 1) = 0$, but $f_X(1)\, f_Y(1) = \dfrac{12}{27} \times \dfrac{6}{27} = \dfrac{8}{81} \neq 0$.

Since $f(1, 1) \neq f_X(1)\, f_Y(1)$, X and Y are **not independent**.

(4) Find $\mathrm{Var}(X - Y)$.

Solution: [Method 2] Since $X \sim B(3, \frac{1}{3})$ and $Y \sim B(3, \frac{2}{3})$,

$$E(X) = 1, \quad E(Y) = 2, \quad \mathrm{Var}(X) = \frac{2}{3}, \quad \mathrm{Var}(Y) = \frac{2}{3}.$$

$$E(XY) = \sum_{x=0}^{3}\sum_{y=0}^{3} xy\, f(x, y) = 1 \times 2 \times \frac{12}{27} + 2 \times 1 \times \frac{6}{27} = \frac{4}{3},$$

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{4}{3} - 2 = -\frac{2}{3}.$$

$$\therefore\ \mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\,\mathrm{Cov}(X, Y) = \frac{2}{3} + \frac{2}{3} - 2\left(-\frac{2}{3}\right) = \frac{8}{3}.$$

[Method 2] Since $Y = 3 - X$,

$$\mathrm{Var}(X - Y) = \mathrm{Var}(2X - 3) = 4\mathrm{Var}(X) = 4np(1 - p) = 4 \times 3 \times \frac{1}{3} \times \frac{2}{3} = \frac{8}{3}.$$

Midterm Exam Solutions

Introduction to Statistics (통계의 이해) **2026 1st Semester**

Year(학년): ___	Section(교반): ___	Cadet Number(교번): _______	Name(성명): _________	Score: _______

4. In a shooting exercise, a cadet aims at a target, and the distance X (cm) from the hit point to the center follows a uniform distribution on $(0, 10)$. The score Y is 10 points if $X < 4$, and 5 points if $4 \leq X < 10$. Answer the following questions. [12 points]

(1) Find the probability density function $f_X(x)$ and the mean $E(X)$.

Solution: Since X follows a uniform distribution on $(0, 10)$, $f_X(x) = \begin{cases} 1/10, & 0 < x < 10, \\ 0, & \text{otherwise.} \end{cases}$ By definition of the mean, $E(X) = \int_0^{10} x \cdot \frac{1}{10}\, dx = 5$.

(2) Find $P(Y = 10)$ and $P(Y = 5)$.

Solution: $P(Y = 10) = P(X < 4) = 4/10$ and $P(Y = 5) = P(4 < X < 10) = 6/10$.

(3) Find the mean $E(Y)$.

Solution: By definition of the mean,

$$E(Y) = \sum_{\forall y} y\, f(y) = 10 \times \frac{4}{10} + 5 \times \frac{6}{10} = 7.$$

5. At a smart farm, the seed size of a crop follows a normal distribution with mean μ and variance σ^2. Let X be the seed size (mm). It is known that $P(X \geq 5) = 0.5$ and $P(X \geq 8.92) = 0.025$. Answer the following questions. [12 points]

(1) Find $E(X)$ and $\mathrm{Var}(X)$.

Solution: Since $P(X \geq 5) = 0.5$ and the normal distribution is symmetric about its mean, $E(X) = \mu = 5$. Since $P(X \geq 8.92) = 0.025$, standardizing X,

$$P\left(\frac{X-5}{\sigma} \geq \frac{8.92-5}{\sigma}\right) = P\left(Z \geq \frac{3.92}{\sigma}\right) = P(Z \geq z_{0.025}) = 0.025.$$

Therefore $\dfrac{3.92}{\sigma} = z_{0.025} = 1.96$, giving $\sigma = \dfrac{3.92}{1.96} = 2$. Thus $\mathrm{Var}(X) = \sigma^2 = 4$.

(2) What is the cutoff for the lowest 2.5% of seed sizes?

Solution: [Method 1] Find x such that $P(X < x) = 0.025$. Standardizing,

$$P\left(\frac{X-5}{2} \leq \frac{x-5}{2}\right) = 0.025 = P(Z \leq -z_{0.025}) = P(Z \leq -1.96).$$

Solving $\dfrac{x-5}{2} = -1.96$ gives $x = 1.08$.

[Method 2] By symmetry of the normal distribution, the arithmetic mean of the 97.5th and 2.5th percentiles equals $\mu = 5$. Solving $\dfrac{x + 8.92}{2} = 5$ gives $x = 1.08$.

6. In a factory, the time to manufacture a type-A chair follows $N(0.7,\ 0.1^2)$ (hours) and the time for a type-B chair follows $N(0.4,\ 0.2^2)$ (hours). Assuming the two manufacturing times are independent, find the probability that manufacturing a type-A chair takes longer than manufacturing a type-B chair. [12 points]

Solution: Let X_A and X_B be the manufacturing times for type-A and type-B chairs, respectively. Then

$$X_A \sim N(0.7,\ 0.1^2), \quad X_B \sim N(0.4,\ 0.2^2).$$

Since X_A and X_B are independent, by the reproductive property of the normal distribution,

$$X_A - X_B \sim N(0.3,\ 0.05).$$

Therefore,

$$P(X_A > X_B) = P(X_A - X_B > 0) = P\left(\frac{(X_A - X_B) - 0.3}{\sqrt{0.05}} > \frac{0 - 0.3}{\sqrt{0.05}}\right)$$
$$\approx P(Z > -1.3416) = 1 - P(Z < -1.3416) = P(Z < 1.3416) = 0.9101.$$

7. It is known that 20% of students at a university live in the campus dormitory. A random sample of $n = 60$ students is selected, and X students among them live in the dormitory. Answer the following questions. [12 points]

(1) Find the approximate distribution of the sample proportion $\hat{p} = X/n$.

Solution: For the sample proportion $\hat{p}$,

$$E(\hat{p}) = 0.2, \quad \text{Var}(\hat{p}) = \frac{0.2 \times 0.8}{60} = 0.0027.$$

Since $n = 60$ is large and $np = 12 > 10$, $n(1-p) = 48 > 10$, by the Central Limit Theorem,

$$\hat{p} \dot{\sim} N(0.2,\ 0.0027).$$

(2) Find $P\left(\hat{p} \geq \frac{1}{6}\right)$.

Solution: Since $\hat{p} \dot{\sim} N(0.2,\ 0.0027)$,

$$P\left(\hat{p} \geq \frac{1}{6}\right) = P\left(\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \geq \frac{1/6 - 1/5}{\sqrt{0.0027}}\right)$$
$$\approx P(Z \geq -0.6416) = P(Z < 0.6416) = 0.7394.$$

Midterm Exam Solutions

Introduction to Statistics (통계의 이해) **2026 1st Semester**

Year(학년): ___	Section(교반): ___	Cadet Number(교번): _______	Name(성명): _________	Score: _______

8. A random sample X_1, X_2, X_3 is drawn from a population with unknown mean μ and known variance σ^2. Consider the following estimators of μ:

$$\hat{\mu}_1 = \frac{X_1 + X_2 + X_3}{3}, \quad \hat{\mu}_2 = \frac{X_1 + 2X_2 + X_3}{4}, \quad \hat{\mu}_3 = X_3.$$

Answer the following questions. [12 points]

(1) Show that $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$ are each unbiased estimators of μ.

Solution:

$$E(\hat{\mu}_1) = \frac{1}{3}(E(X_1) + E(X_2) + E(X_3)) = \frac{1}{3}(\mu + \mu + \mu) = \mu,$$

$$E(\hat{\mu}_2) = \frac{1}{4}(E(X_1) + 2E(X_2) + E(X_3)) = \frac{1}{4}(\mu + 2\mu + \mu) = \mu,$$

$$E(\hat{\mu}_3) = E(X_3) = \mu.$$

Therefore $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$ are all unbiased estimators of μ.

(2) Find the most efficient estimator among $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$. (Hint: $\mathrm{Var}(\hat{\mu}_1) = \sigma^2/3$.)

Solution: Since all three estimators are unbiased, we compare their variances:

$$\mathrm{Var}(\hat{\mu}_1) = \frac{\sigma^2}{3},$$

$$\mathrm{Var}(\hat{\mu}_2) = \frac{1}{16}\big(\mathrm{Var}(X_1) + 4\,\mathrm{Var}(X_2) + \mathrm{Var}(X_3)\big) = \frac{1}{16}(\sigma^2 + 4\sigma^2 + \sigma^2) = \frac{6\sigma^2}{16} = \frac{3\sigma^2}{8},$$

$$\mathrm{Var}(\hat{\mu}_3) = \sigma^2.$$

Since $\mathrm{Var}(\hat{\mu}_1) < \mathrm{Var}(\hat{\mu}_2) < \mathrm{Var}(\hat{\mu}_3)$, the most efficient estimator of μ is $\hat{\mu}_1$.

9. The weight (g) of soap bars produced at a company follows a normal distribution. A random sample of 25 bars was selected and the sample mean was $\bar{x} = 97$. Answer the following questions. [12 points]

(1) If the population standard deviation is $\sigma = 2$, find a 95% confidence interval for the population mean μ.

Solution: When σ^2 is known, a $100(1-\alpha)\%$ confidence interval for μ from a normal population is

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

With $\bar{x} = 97$, $\sigma = 2$, $n = 25$, and $z_{0.025} = 1.96$, a 95% confidence interval for μ is

$$\left[97 - 1.96 \cdot \frac{2}{\sqrt{25}},\ 97 + 1.96 \cdot \frac{2}{\sqrt{25}}\right] \approx [96.2160,\ 97.7840].$$

(2) If the sample standard deviation is $s = 2.0616$, find a 90% confidence interval for the population mean μ.

Solution: When σ^2 is unknown, a $100(1-\alpha)\%$ confidence interval for μ is

$$\left[\bar{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}},\ \bar{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right].$$

With $\bar{x} = 97$, $s = 2.0616$, $n = 25$, and $t_{0.05}(24) = 1.711$, a 90% confidence interval for μ is

$$\left[97 - 1.711 \cdot \frac{2.0616}{\sqrt{25}},\ 97 + 1.711 \cdot \frac{2.0616}{\sqrt{25}}\right] \approx [96.2945,\ 97.7055].$$

※ Reference Table 1

probability mass function	probability density function	statistic	central limit theorem	standard normal
confidence interval	efficiency	estimate	estimator	t
statistic	probability distribution	uniform	parameter	population mean
unbiased	population variance	efficient	normal	sample variance
alternative hypothesis	test statistic	rejection region	significance level	null hypothesis
type I error	type II error	p-value	power	right-sided test
population proportion	consistent	chi-squared	binomial	$n+1$
n	$n-1$	$n-2$	cumulative distribution function	joint probability mass function

※ Reference Table 2 Let $Z \sim N(0,1)$.

$P(Z<0.005)=0.502$	$P(Z<0.01)=0.504$	$P(Z<0.025)=0.51$	$P(Z<0.05)=0.5199$
$P(Z<0.2425)=0.5958$	$P(Z<1)=0.8414$	$P(Z<1.3416)=0.9101$	$P(Z<1.5)=0.9332$
$P(Z<0.6416)=0.7394$	$P(Z<2)=0.9772$	$P(Z<2.2361)=0.9873$	$P(Z<2.451)=0.9929$
$P(Z<2.5)=0.9938$	$P(Z<2.5454)=0.9945$	$P(Z<2.666)=0.9962$	$z_{0.0007}=3.21$
$z_{0.005}=2.5759$	$z_{0.01}=2.3263$	$z_{0.0228}=1.999$	$z_{0.025}=1.96$
$z_{0.0328}=1.8417$	$z_{0.05}=1.645$	$z_{0.0654}=1.511$	$z_{0.0668}=1.5$
$z_{0.0787}=1.4142$	$z_{0.1587}=1$	$z_{0.2389}=0.71$	$z_{0.2743}=0.6$
$z_{0.3088}=0.4992$	$t_{0.0034}(8)=3.6134$	$t_{0.0037}(8)=3.5675$	$t_{0.0043}(8)=3.4586$
$t_{0.005}(8)=3.3554$	$t_{0.0060}(8)=3.2346$	$t_{0.01}(8)=2.8965$	$t_{0.025}(8)=2.3060$
$t_{0.05}(8)=1.8595$	$t_{0.0024}(10)=3.6134$	$t_{0.0026}(10)=3.5675$	$t_{0.0031}(10)=3.4586$
$t_{0.0045}(10)=3.2346$	$t_{0.025}(18)=2.101$	$t_{0.05}(18)=1.734$	$t_{0.025}(19)=2.093$
$t_{0.05}(19)=1.729$	$t_{0.005}(20)=2.8453$	$t_{0.025}(20)=2.5280$	$t_{0.025}(20)=2.0860$
$t_{0.0371}(20)=1.8839$	$t_{0.05}(20)=1.7247$	$t_{0.1069}(20)=1.2839$	$t_{0.2075}(20)=0.8323$
$t_{0.2829}(20)=0.5839$	$t_{0.005}(24)=2.797$	$t_{0.025}(24)=2.064$	$t_{0.05}(24)=1.711$
$t_{0.1}(24)=1.318$			