

Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

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

1. Read each statement below. Using the word box in the Reference Materials, fill in each blank with the appropriate term, or determine whether the statement is **True (O)** or **False (X)**. [8 points]

- (1) A numerical value describing a characteristic of the population is called a (_____).
- (2) Suppose a random sample $X_1, \dots, X_n$ of size n is taken from a Bernoulli distribution $B(1, p)$ with probability of success p . A point estimate of the population proportion p is (_____), $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$
- (3) The rejection region is the set of test statistic values for which H_0 is rejected with a certain probability α . This probability α is called the (_____).
- (4) The (_____) of the i -th observation (x_i, y_i) is the difference between the observed y_i and the predicted $\hat{y}_i$.
- (5) The central limit theorem states that, regardless of the population distribution, the distribution of the sample mean $\bar{X}$ approaches a normal distribution as the sample size n increases. (O / X)
- (6) If $X_1, \dots, X_n$ is a random sample from a normal population $N(\mu, \sigma^2)$ and S^2 is the sample variance, then $V = \frac{(n-1)S^2}{\sigma^2}$ follows a chi-squared distribution with n degrees of freedom. (O / X)
- (7) For two estimators $\hat{\theta}_1, \hat{\theta}_2$ of a parameter θ , if $\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2)$, then $\hat{\theta}_1$ is said to be more efficient than $\hat{\theta}_2$. (O / X)
- (8) Least squares regression tries to maximize the sum of squared errors. (O / X)

2. A random sample $X_1, X_2, \dots, X_n$ of size n is drawn from a population with probability density function:

$$f(x) = \begin{cases} \frac{1}{2\theta}, & 0 < x < 2\theta, \\ 0, & \text{otherwise,} \end{cases}$$

with $\theta > 0$. Answer the following. [10 points]

(You may use $E(X_1) = \theta$ and $\text{Var}(X_1) = \frac{\theta^2}{3}$)

(1) Show that the sample mean $\bar{X}$ is an unbiased estimator of θ .

(2) Show that the sample mean $\bar{X}$ is a consistent estimator of θ .

Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

3. A buyer wants to test whether the average diameter of nuts from a supplier is different from $\mu_0 = 12.7$ mm. A random sample of $n = 10$ nuts is selected and their diameters $X_1, X_2, \dots, X_{10}$ are measured. Test at significance level $\alpha = 0.05$. Assume the nut diameters follow a normal distribution. [12 points]

(1) State the null and alternative hypotheses.

(2) Find the test statistic and its null distribution.

(3) The measured diameters of the 10 selected nuts are shown below. Find the observed value of the test statistic.

Nut (i)	1	2	3	...	8	9	10
Diameter (x_i , mm)	12.68	12.79	12.85	...	12.67	12.79	12.65

$$\ast \bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = 12.744, \quad s = \sqrt{\frac{1}{10-1} \sum_{i=1}^{10} (x_i - \bar{x})^2} = 0.0783$$

(4) Using the result in (3), report the p -value and complete the hypothesis test. State the conclusion in the context of the data.

Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

4. The average recovery time for an existing health supplement is $\mu_0 = 20$ days. A team claims that their new supplement gives a smaller average recovery time (less than 20 days). To test this, $n = 16$ people take the new supplement and their recovery times $X_1, X_2, \dots, X_{16}$ are measured. Test at significance level $\alpha = 0.05$. Assume the recovery time follows a normal distribution with standard deviation $\sigma = 4$ days. [12 points]

(1) State the null and alternative hypotheses.

(2) Find the test statistic and its null distribution.

(3) Find the rejection region at significance level $\alpha = 0.05$.

(4) The sample mean recovery time is $\bar{x} = 18.2$ days. Find the observed value of the test statistic.

(5) Using the result in (4) and the rejection region, complete the hypothesis test. State the conclusion in the context of the data.

Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

5. A company claims that its new earplugs reduce noise by more than $\delta_0 = 2$ dB on average compared with the existing earplugs. To test this, $n_A = 40$ new and $n_B = 50$ existing earplugs are randomly selected, and their noise reductions $X_{A1}, \dots, X_{A40}$ and $X_{B1}, \dots, X_{B50}$ are measured. Test at significance level $\alpha = 0.01$. Assume the noise reductions of the new and existing earplugs follow normal distributions with means μ_A, μ_B and standard deviations $\sigma_A = 0.5, \sigma_B = 0.6$. [15 points]

(1) State the null and alternative hypotheses.

(2) $\bar{X}_A - \bar{X}_B$ is used as the test statistic for (1), where $\bar{X}_A$ and $\bar{X}_B$ are the sample means for the new and existing earplugs, respectively. Find the null distribution of this test statistic.

(3) Find the rejection region at significance level $\alpha = 0.01$.

(4) The sample mean noise reductions of the selected new and existing earplugs are $\bar{x}_A = 27.48$ and $\bar{x}_B = 25.17$. Find the observed value of the test statistic and test the hypothesis at significance level $\alpha = 0.01$. State the conclusion in the context of the data.

(5) When the true difference of the two population means is $\mu_A - \mu_B = 2.5$, find the power of the test.

Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

6. A newly developed protein supplement is tested to see whether it increases grip strength (kg). For $n = 12$ randomly selected people, grip strength is measured before, X_{Ai} , and after, X_{Bi} ($i = 1, \dots, 12$), taking the supplement. Using a paired test at significance level $\alpha = 0.05$, we test whether the average grip strength is greater after than before. Assume the before, X_A , and after, X_B , grip strengths each follow a normal distribution with means μ_A and μ_B , and define $\mu_{\text{diff}} = \mu_A - \mu_B$. The hypotheses are $H_0 : \mu_{\text{diff}} = 0$ and $H_A : \mu_{\text{diff}} < 0$. [14 points]

(1) Find the test statistic and its null distribution.

(2) Find the rejection region at significance level $\alpha = 0.05$.

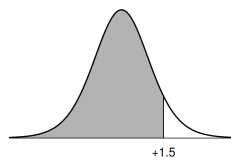
(3) The partial R output for this problem is shown below. Using the partial R output and the rejection region in (2), complete the hypothesis test. State the conclusion in the context of the data.

Paired t-test

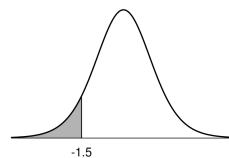
```
data: before and after
t = -1.5, df = _____, p-value = _____
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
 -Inf 0.5917717
sample estimates:
mean difference
      -3
```

(4) Among the figures below, choose the one whose shaded (gray) region best corresponds to the p -value in (3).

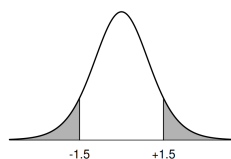
①



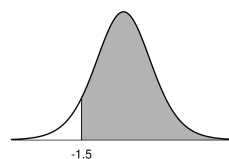
②



③



④



Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

7. A café headquarters wants to test whether there is a correlation (ρ) between the floor area (X) and the weekly sales (Y) of a branch. A random sample of $n = 10$ branches gives the results below. Test at significance level $\alpha = 0.05$. Assume (X, Y) follows a bivariate normal distribution. [14 points]

Floor area (x_i , pyeong)	14	19	25	...	26	17	20
Weekly sales (y_i , 10,000 KRW)	245	275	378	...	390	268	292

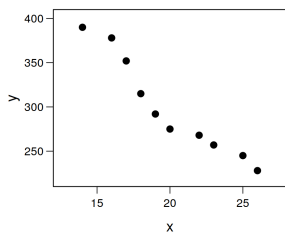
- (1) State the null and alternative hypotheses.

- (2) The test statistic for (1) is $\frac{R\sqrt{n-2}}{\sqrt{1-R^2}}$, where R is the sample correlation coefficient. Find the null distribution of this test statistic.

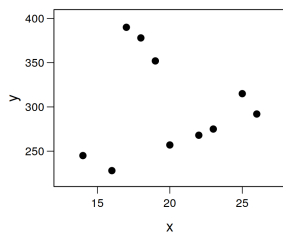
- (3) Find the rejection region at significance level $\alpha = 0.05$.

- (4) When the observed sample correlation coefficient is $r = 0.9652$, choose the figure below that best represents the scatter plot of the floor area (X) and the weekly sales (Y).

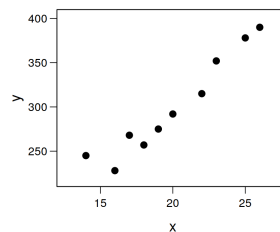
①



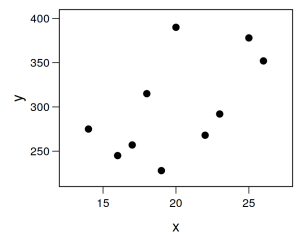
②



③



④



- (5) When the observed sample correlation coefficient is $r = 0.9652$, find the observed value of the test statistic and test the hypothesis using the rejection region in (3). State the conclusion in the context of the data.

Final Exam

Introduction to Statistics(통계의 이해 영어강의).

2026 1st semester

Section(교반): _____ Cadet Number(교번): _____ Name(성명): _____ Score: _____

8. A shooting instructor analyzes the relationship between a recruit's practice time (x , hours) and shooting-test score (Y , points) using the simple linear regression model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma^2)$. A random sample of $n = 10$ recruits is analyzed using the R function `lm()`, with the results shown below. Answer the following. [15 points]

```
> x <- c(10, 12, 14, ..., 24, 26, 28)
> y <- c(64, 80, 60, ..., 94, 80, 97)
> summary(lm(y ~ x))
```

```
Call:
lm(formula = y ~ x)

Residuals:
    Min       1Q   Median       3Q      Max
-12.030  -8.614  -1.500   8.462  10.758

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  52.5152    11.0408   4.756  0.00143 **
x             1.3939     0.5562   2.506  0.03659 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.1 on 8 degrees of freedom
Multiple R-squared:  0.4398,    Adjusted R-squared:  0.3698
F-statistic:  6.28 on 1 and 8 DF,  p-value: 0.03659
```

(1) Using the R output, find the values in the table below.

Estimated regression coefficients		Estimated regression equation
$\hat{\beta}_0$	$\hat{\beta}_1$	

(2) The first recruit's practice time is $x_1 = 10$ and shooting-test score is $y_1 = 64$. Using the simple linear regression analysis, find the fitted value $\hat{y}_1$ and the residual e_1 .

(3) Suppose $\hat{\sigma}^2 = MSE = 102.01$. Find $\sum_{i=1}^{10} e_i^2$ for the residuals $e_1, \dots, e_{10}$.

(4) Using the R output, we test the hypothesis $H_0 : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$ at significance level $\alpha = 0.05$ to see whether the practice time affects the shooting-test score.

a. The observed value of the test statistic is (_____), and its p -value is (_____). Therefore, at significance level $\alpha = 0.05$, we (_____) H_0 .

b. State the conclusion in the context of the data.