실험계획법 실습

인하대학교 대학원
홍규성

- **Basic Statistical Concepts**
- **Sampling and Sampling Distributions**
- **Inference about the difference in Mean, Randomized Designs**
- **Inference about the difference in Means, Paired Comparison Designs**
- **Inference about the variance of Normal Distribution**

◆ Basic Statistical Concepts

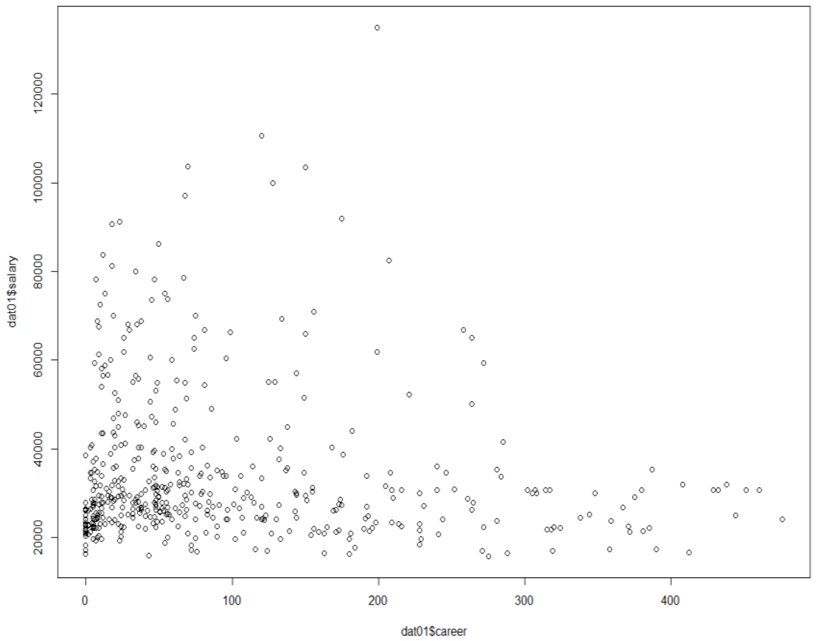
1. Graphical Description of Variability
   - Plot
   - Bar chart
   - Box plot
   - Histogram

2. Probability Distributions
   - Mean
   - Variance
   - Standard Deviation

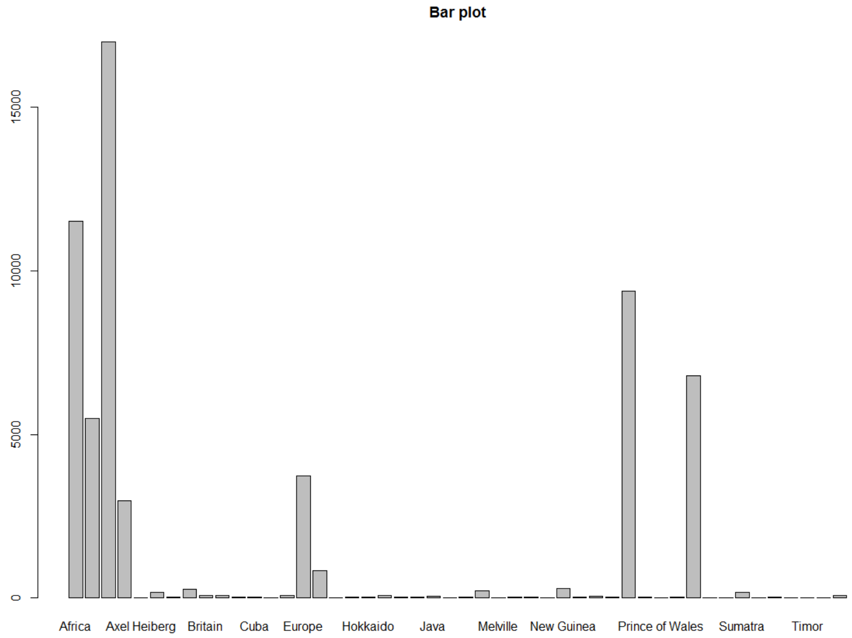
◆ Basic Statistical Concepts

1. Graphical Description of Variability

- Scatter Plot

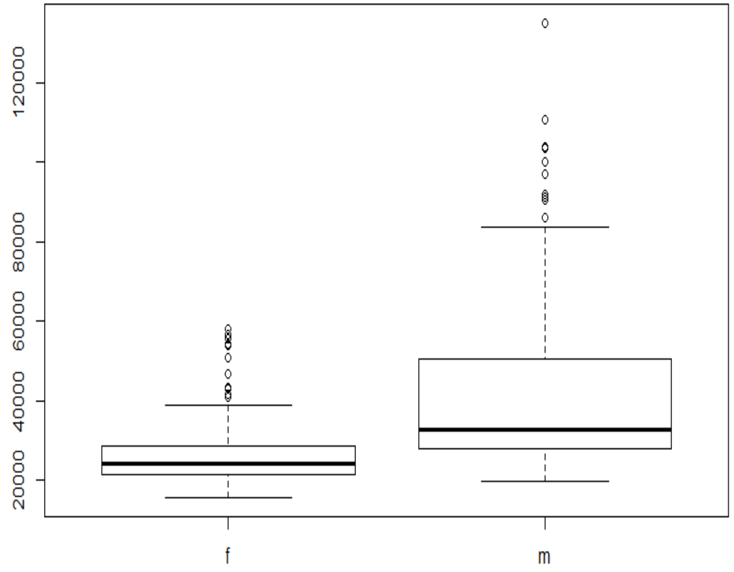


- Bar chart

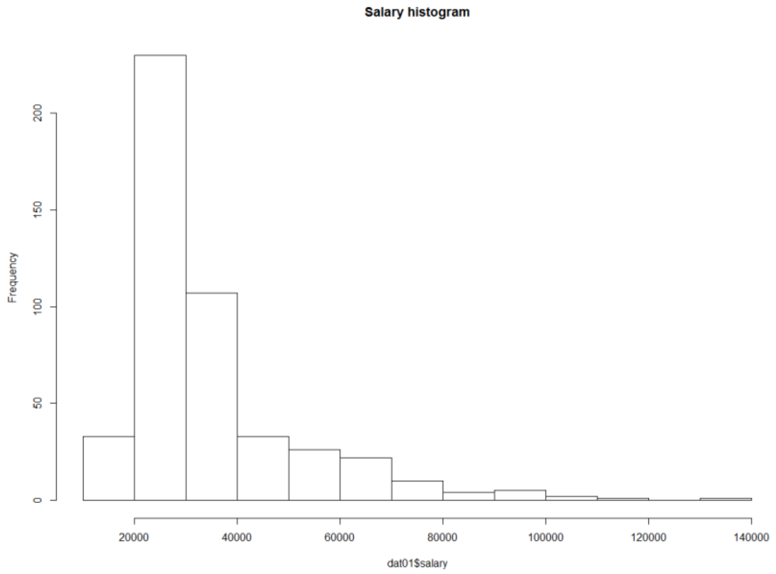


Basic Statistical Concepts

- **Box plot**



- **Histogram**



◆ Sampling and Sampling Distributions

1. Statistics: Point Estimation

	Parameters		Estimators	(Sample) Estimates
Mean:	$\mu_Y = E[Y]$	←	$\overline{Y} = \sum Y_j / n$	$\overline{y} = \sum y_j / n$
Variance:	$\sigma_Y^2 = E[(Y-\mu_Y)^2]$	←	$S^2 = \dfrac{\sum (Y_i - \overline{Y})^2}{n-1}$	$s^2 = \dfrac{\sum (y_i - \overline{y})^2}{n-1}$
Standard Deviation:	σ_Y	←	S	s

◆ Sampling and Sampling Distributions

1. Statistics: Point Estimation

1) $E(c) = c$
2) $E(y) = \mu$
3) $E(cy) = cE(y) = c\mu$
4) $V(c) = 0$
5) $V(y) = \sigma^2$
6) $V(cy) = c^2V(y) = c^2\sigma^2$
7) $E(y_1 + y_2) = E(y_1) + E(y_2) = \mu_1 + \mu_2$
8) $V(y_1 + y_2) = V(y_1) + V(y_2) + 2Cov(y_1, y_2)$

◆ Sampling and Sampling Distributions

2. The Central Limit Theorem

If $y_1, y_2\cdots, y_n$ is a sequence of n independent and identically distributed random variables with $E(y_i) = \mu$ and $V(y_i) = \sigma^2$ (both finite) and x= $y_1+y_2+\cdots+y_n$, then the limiting form of the distribution of

$$Z_n = \frac{x-n\mu}{\sqrt{n\sigma^2}}$$

as n → ∞ , is the standard normal distribution.

◆ Inference about the difference in Mean, Randomized Designs

1. Hypotheses Testing

- **Hypothesis: Null(H_0) versus Alternative(H_1)**

  Determine whether null is rejected or not rejected

- **Critical Regions:** CR

  Reject null hypothesis if the test statistics is in the critical region

- **Type I and Type II Errors**

  Type I Error: Reject H_0 when H_0 is true

  → $\alpha = P[\text{type I error}] = P[\text{reject } H_0 \setminus H_0 \text{ true}]$

  Type II Error: Not Reject H_0 when H_0 is not true ($\theta \notin H_0$)

  → $\beta(\mu) = P[\text{type II error}] = P[\text{not reject } H_0 \setminus \mu \notin H_0 \text{ true}]$

◆ Inference about the difference in Mean, Randomized Designs

1. Hypotheses Testing

- **Testing**

i) Classical Approach

Choose α and determine CR. If test statistic is in CR, reject H_0

ii) $p-value$ Approach

Compute $p-value$ of the statistic. If $p <$ (some α), reject H_0

← The probability of obtaining a value of test statistic that is at least as extreme as the calculated value when H_0 is true

◆ Inference about the difference in Mean, Randomized Designs

2. The Case Where $\sigma_1^2 \neq \sigma_2^2$

Independent Samples and (σ_1^2, σ_2^2) are unknown but unequal

$H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$ $(H_1: \mu_1 > \mu_2)$

- Test Statistic: $t' = \dfrac{\overline{Y}_1 - \overline{Y}_2}{\sqrt{(S^2/n_1) + (S^2/n_2)}} \sim t_{df}$, under H_0

  (Normal (or large Non-Normal) with $df = \dfrac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1+1} + \dfrac{(s_2^2/n_2)^2}{n_2+1}} - 2$)

- $100(1-\alpha)$ C. I. for $\mu_1 - \mu_2$:

$$(\overline{y}_1 - \overline{y}_2) \pm t_{1-\alpha/2} \times \sqrt{(s_1^2/n_1) + (s_2^2/n_2)}$$

◆ Inference about the difference in Mean, Randomized Designs

3. The Case Where σ_1^2 and σ_2^2 Are Known

Independent Samples and (σ_1^2, σ_2^2) are known

$H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$ $(H_1 : \mu_1 > \mu_2)$

- Test Statistics: $Z = \dfrac{\overline{Y}_1 - \overline{Y}_2}{\sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}} \sim N(0,1)$, under H_0

  (Normal Assumptions, or Not Normal with $n_1, n_2 \geq 25$)
- Decision Rule: H_0 if $|z| \geq Z_{1-\alpha/2}$ $(z \geq Z_{1-\alpha})$
- $100(1-\alpha)$ C. I. for $\mu_1 - \mu_2$: $(\overline{y}_1 - \overline{y}_2) \pm Z_{1-\alpha/2} \times \sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}$

◆ Inference about the difference in Mean, Randomized Designs

4. Comparing a Single Mean to a Specified Value

$H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ $(H_1 : \mu > \mu_0)$

Test Procedure:

i) σ is known

- Test Statistics: $Z = \dfrac{\overline{Y} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$, under H_0 (Normal Assumption)
- Decision Rule: H_0 if $z \geq Z_{1-\alpha/2}$ or $z \leq Z_{\alpha/2}$ $(z \geq Z_{1-\alpha})$
- $100(1-\alpha)$ C. I. for μ : $\overline{y} \pm Z_{1-\alpha/2} \times \sigma / \sqrt{n}$

ii) σ is unknown

- Test Statistics: $t = \dfrac{\overline{Y} - \mu_0}{S / \sqrt{n}} \sim t_{n-1}$, under H_0

  (Normal Assumption, or Not Normal but large n (≥ 30))
- Decision Rule: H_0 if $|t| \geq t_{1-\alpha/2}$ $(t \geq t_{1-\alpha})$
- $100(1-\alpha)$ C. I. for μ : $\overline{y} \pm t_{1-\alpha/2} \times s / \sqrt{n}$

◆ Inference about the difference in Mean, Paired Comparison Designs

1. The Paired Comparison Problem

- Differenced Data: $D = Y_1 - Y_2$ (data: $d_i = y_{1i} - y_{2i}$)
- Hypothesis: $H_0 : \mu_D = 0$ versus $H_1 : \mu_D \neq 0$ $(H_1 : \mu_D > 0)$
- Test Statistics: $t = \dfrac{\overline{D}}{S_D / \sqrt{n}} \sim t_{n-1}$, under H_0

  (Differenced Population is Normal, or Not Normal but large n (≥ 30))
- Decision Rule: H_0 if $|t| \geq t_{1-\alpha/2}$ $(t \geq t_{1-\alpha})$
- $100(1-\alpha)$ C. I. for $\mu_D = \mu_1 - \mu_2$: $\overline{d} \pm t_{1-\alpha/2} \times s_D / \sqrt{n}$

◆ Inference about the Variances of Normal Distributions

1. Test on a Single Variance

- $H_0 : \sigma^2 = \sigma_0^2$ versus $H_1 : \sigma^2 \neq \sigma_0^2$ $(H_1 : \sigma^2 > \sigma_0^2)$
- Test Statistic: $W = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$, under H_0 (with Normal Assumption)
- Decision Rule: Reject H_0 if $w \geq \chi^2_{1-\alpha/2}$ or $w \leq \chi^2_{\alpha/2}$ $(w \geq \chi^2_{1-\alpha})$
- $100(1-\alpha)$ Confidence Interval for σ^2:

  $[(n-1)s^2 / \chi^2_{1-\alpha/2},\ (n-1)s^2 / \chi^2_{\alpha/2}]$

◆ Inference about the Variances of Normal Distributions

2. Test on Two Variance : Independent Samples

- $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 \neq \sigma_2^2$ $(H_1 : \sigma_1^2 > \sigma_2^2)$
- Test Statistic: $F = \dfrac{S_1^2}{S_2^2} \sim F_{n_1 - 1, n_2 - 1}$, under H_0

  (with independent Normal Assumptions)
- Decision Rule: Reject H_0 if $f \geq F_{1-\alpha/2}$ or $f \leq F_{\alpha/2}$ $(f \geq F_{1-\alpha})$
- $100(1-\alpha)$ C. I. for σ^2: $[(s_1^2 / s_2^2) \times F_{\alpha/2},\ (s_1^2 / s_2^2) \times F_{1-\alpha/2}]$

감사합니다.
See You