# [Chapter 8. Estimation]

# 8.1 Introduction

- Purpose of Statistics, estimation and testing

  · make inference about the population by using the information contained in a sample taken from the population of our interest

  · find a *statistic* for a *unknown target parameter* characterizing the population and the *sampling distribution* of the statistic in making statistical inferences

  : (point/interval) estimation and testing

- Estimation and Estimator

  · Estimation in two different forms

  i) Point estimation, ii) Interval estimation

  (example) interested in estimating the unknown mean waiting time $\mu$ at a supermarket checkout station

  i) point estimation by a single number : an expert considers 10 minutes as the estimate of $\mu$.

  ii) interval estimation by the two values enclosing $\mu$ : $\mu$ will fall in between 6 and 14(i.e., [6, 14])

· (Definition 8.1) An estimator is a formula that tells how to calculate the value of an (point/interval) estimate based on the measurements contained in a sample.

(example) Firing a revolver at a red target

: one shot does not tell us if he/she is an expert, but many shots(say 100) shots might provide sufficient amount of evidence

: (estimator - revolver), (estimate - a single shot), and (parameter of interest - a red target)

(example) Suppose one is interested in the center (i.e., location parameter) of the population, $\mu$. Given $n$ random variables, $Y_1, \ldots, Y_n$ from the population, the sample mean is $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$(which is also a random variable). After observing a particular value of a random variable, $Y_1 = y_1, \ldots, Y_n = y_n$, then $\bar{Y} = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

(example) Suppose we are interested in the mean of UMass female students' height, $\mu$. So, we measure heights of 100 female students and calculate sample mean, $\bar{y}$ using a sample mean formula $\bar{Y}$.

: but we can not evaluate the goodness of point estimation procedure based on one single constant value, $\bar{y}$. We would evaluate this point estimation procedure after this procedure is used many times.

: Thus, we do repeated sampling(i.e., obtain 1000 samples of size 100 from the population) and calculate sample mean from each sample, $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_{1000}$.

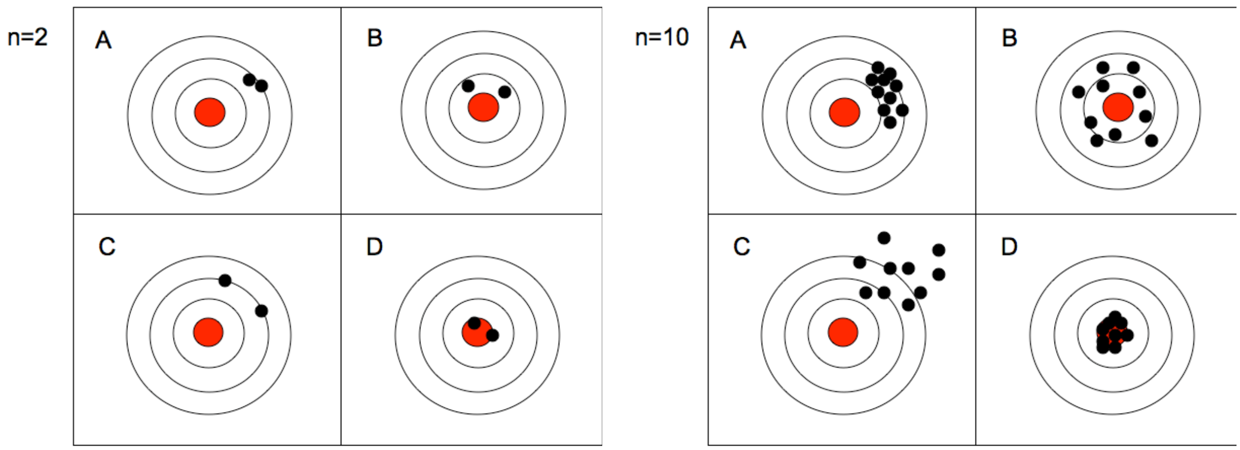
: Finally we construct a frequency distribution of 1000 sample means and see how closely the distribution clusters around the true mean of UMass female students' height. Note a frequency distribution of 1000 sample means is an approximation to the sampling distribution of $\bar{Y}$.

· Many different estimators for the same population parameter. Then we need to know how we can find better estimators

(example) $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, sample mean is one possible point estimator of $\mu$, the center (i.e., location parameter) of the population. How about the median or mode of $n$ random variables?

## 8.2 Bias and Mean Square Error of Point Estimators

Four students(A,B, C, D) fire a revolver at a red target $n$ times.




- Variation among $n$ shots

  : $\quad < \quad < \quad <$

  :

- Average distance between $n$ shots and the target

  : $\quad < \quad < \quad < \quad <$

  :

Suppose $Y_1, \ldots, Y_n$ constitute a random sample from a population with a parameter $\theta$ of our interest. Let $\hat{\theta} = \hat{\theta}(Y_1, \ldots, Y_n)$(random quantity) be a point estimator for a parameter $\theta$.

- Since $\hat{\theta}$ is a statistic for $\theta$, it has its own sampling distribution, say $f(\hat{\theta})$.

- Unbiasedness

  (Def 8.2 and 8.3) The bias of $\hat{\theta}$ is given by $B(\hat{\theta}) = E(\hat{\theta}) - \theta$. If $B(\hat{\theta}) = 0$, $\hat{\theta}$ is an unbiased estimator of $\theta$. Otherwise, $\hat{\theta}$ is a biased estimator.

- Variance and standard error

  (Def) The variance of the sampling distribution of $\hat{\theta}$ is given by $\sigma^2_{\hat{\theta}} = V(\hat{\theta}) = E(\hat{\theta} - E(\hat{\theta}))^2$. The standard deviation of the sampling distribution of $\hat{\theta}$ is given by $\sigma_{\hat{\theta}} = \sqrt{\sigma^2_{\hat{\theta}}} = \sqrt{V(\hat{\theta})}$. We call $\sigma_{\hat{\theta}}$ the standard error of $\hat{\theta}$.

- Mean square error(MSE)

  (Def 8.4) The MSE of $\hat{\theta}$ is given by $MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = V(\hat{\theta}) + \left(B(\hat{\theta})^2\right)$.

(Exercise) Suppose that $Y_1$, $Y_2$, $Y_3$ denote a random sample from an exponential distribution with a parameter $\theta$. Consider the following four estimators of $\theta$ : $\widehat{\theta}_1 = Y_1$, $\widehat{\theta}_2 = (Y_1 + Y_2)/2$, $\widehat{\theta}_3 = (Y_1 + 2Y_2)/3$ and $\theta_4 = \bar{Y}$.

a. Which of these estimators are unbiased?

b. Among the unbiased estimators, which has the smallest variance?

(Exercise) Suppose $Y$ has a binomial distribution with parameters $n$ and $p$(i.e., $Y \sim b(n, p)$).

a. Show that $\widehat{p}_1 = Y/n$ is an unbiased estimator of $p$.

b. Consider another estimator, $\widehat{p}_2 = (Y + 1)/(n + 2)$. Then find the bias of $\widehat{p}_2$.

c. Derive $MSE(\widehat{p}_1)$ and $MSE(\widehat{p}_2)$.

## 8.3 Some Common Unbiased Point Estimators

- Methods for point estimators : Chapter 9.
- Common unbiased $\hat{\theta}$ (Table 8.1)

| | $\theta$ | Sample size | $\hat{\theta}$ | $E(\hat{\theta})$ | $\sigma_{\hat{\theta}}$ | $\hat{\sigma}_{\hat{\theta}}$ |
|---|---|---|---|---|---|---|
| 1) | $\mu$ | $n$ | $\bar{Y}$ | $\mu$ | $\sigma/\sqrt{n}$ | $S/\sqrt{n}$ |
| 2) | $p$ | $n$ | $\frac{Y}{n}$ | $p$ | $\sqrt{\frac{pq}{n}}$ | $\sqrt{\frac{\frac{Y}{n}(1-\frac{Y}{n})}{n}}$ |
| 3) | $\mu_1 - \mu_2$ | $n_1$,$n_2$ | $\bar{Y}_1 - \bar{Y}_2$ | $\mu_1 - \mu_2$ | $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ | $\sqrt{\frac{S_1^2}{n_1} + \frac{S_1^2}{n_2}}$ |
| 4) | $p_1 - p_2$ | $n_1$,$n_2$ | $\frac{Y_1}{n_1} - \frac{Y_2}{n_2}$ | $p_1 - p_2$ | $\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$ | $\sqrt{\frac{\frac{Y_1}{n_1}(1-\frac{Y_1}{n_1})}{n_1} + \frac{\frac{Y_2}{n_2}(1-\frac{Y_2}{n_2})}{n_2}}$ |

1) For $n$ random samples, $Y_1, \ldots, Y_n$ with $E(Y_i) = \mu$(population mean) and $V(Y_i) = \sigma^2$, $\bar{Y}$ for $\mu$

2) For $Y \sim b(n, p)$ and $q = 1 - p$, sample proportion $\frac{Y}{n}$ for $p$

3) For $n_1$ random samples $Y_{11}, \ldots, Y_{1n_1}$ with $E(Y_{1i}) = \mu_1$ and $V(Y_{1i}) = \sigma_1^2$, and $n_2$ random samples $Y_{21}, \ldots, Y_{2n_2}$ with $E(Y_{2j}) = \mu_2$ and $V(Y_{2j}) = \sigma_2^2$ where $i = 1, \ldots, n_1$ and $j = 1, \ldots, n_2$, $\bar{Y}_1 - \bar{Y}_2$ for $\mu_1 - \mu_2$.

4) For $Y_1 \sim b(n, p_1)$ $(q_1 = 1 - p_1)$ and $Y_2 \sim b(n, p_2)(q_2 = 1 - p_2)$ , difference in the sample proportions $\frac{Y_1}{n_1} - \frac{Y_2}{n_2}$ for $p_1 - p_2$.

(note) the two samples from two populations in 3) and 4) are independent.

- Comments for estimators in Table 8.1

  $\cdot$ 1) and 3) have valid $E(\widehat{\theta})$ and $\sigma_{\widehat{\theta}}$ regardless of the form of the population distribution, $p(y)$ or $f(y)$

  $\cdot$ 1) - 4) have sampling distributions, either $p(\widehat{\theta})$ or $f(\widehat{\theta})$ that are approximately normal for large samples (by C.L.T in Chapter 7.3)

  $\cdot$ 1) - 4) are unbiased with near-normal(or bell-shaped) sampling distributions for moderate-sized samples

- For $n$ random samples $Y_1, \ldots, Y_n$ with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$, an estimator for the population variance, $\sigma^2$ is (see example 8.1)

  $\cdot$ $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ : unbiased

  $\cdot$ $S'^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ : biased

- Goodness of a point estimator : how much faith can one place in the validity of statistical inference? (See Chapter 8.4)

## 8.4 Evaluating the Goodness of a Point Estimator

- Reasonable measure of the goodness of $\widehat{\theta}$ for $\theta$ :

  (Def 8.5) The error of estimation $\epsilon$ is $\epsilon = \mid \widehat{\theta} - \theta \mid$.

  · hopes $\epsilon$ to be as small as possible

  · varies randomly in repeated sampling

  · $\epsilon$ is a random quantity, as $\widehat{\theta} = \widehat{\theta}(Y_1, \ldots, Y_n)$ is a also random variable.

- Probability statements about $\epsilon$ : $P(\epsilon < b)$

  · Suppose $\widehat{\theta}$ satisfies the following properties, i) $E(\widehat{\theta}) = \theta$ and ii) its sampling distribution is symmetric at $\theta$.

  Then, $P(\epsilon = \mid \widehat{\theta} - \theta \mid < b) = P(\theta - b < \widehat{\theta} < \theta + b) \approx$ the fraction of times, **in repeated sampling**, that $\widehat{\theta}$ falls within $b$ units of $\theta$ for small $b$(probabilistic bound on $\epsilon$).

- Calculation of $b$ so that $P(\epsilon < b) = .90$

  $\cdot$ If we know the probability distribution, $f(\widehat{\theta})$ of $\widehat{\theta}$, $b$ satisfies $\int_{\theta-b}^{\theta+b} f(\widehat{\theta})d\widehat{\theta} = .90$.

  $\cdot$ If we do not know $f(\widehat{\theta})$, can we obtain $b$?

  i) use Tchebysheff's theorem (p.146 or p.207) to obtain an approximate bound on $\epsilon$.

  : Suppose $\widehat{\theta}$ is unbiased. Then $P(\epsilon = \mid \widehat{\theta} - \theta \mid < b = k\sigma_{\widehat{\theta}}) \geq 1 - 1/k^2$ for $k \geq 1$.

  : for $k = 2$, $P(\epsilon < b = 2\sigma_{\widehat{\theta}}) \geq .75$, which is very conservative.

  ii) use a 2-standard error bound, $b = 2\sigma_{\widehat{\theta}}$ or $2\widehat{\sigma}_{\widehat{\theta}}$

  : $P(\epsilon = \mid \widehat{\theta} - \theta \mid < b = 2\sigma_{\widehat{\theta}})$ is near .95 in many situations (Table 8.2).

(Example 8.2)

(Example 8.3)

# 8.5 Confidence interval (Interval estimator)

- Interval estimator, $[\widehat{\theta}_L, \widehat{\theta}_U]$ : a procedure calculating an interval of probable values of an unknown population parameter, $\theta$ by using the sample measurements

  · indicate the reliability of an estimate, as it represents a range of values around an estimate that include $\theta$(with a certain probability in repeated sampling)

  · its length and location are random quantities, as one or both of endpoints in the interval vary randomly from sample to sample (i.e., $\widehat{\theta}_L = \widehat{\theta}_L(Y_1, \ldots, Y_n), \widehat{\theta}_U = \widehat{\theta}_U(Y_1, \ldots, Y_n)$)

  · $[\widehat{\theta}_L, \widehat{\theta}_U]$ are written with a percentage; what does this percentage represent?

- Confidence of the interval, $100(1-\alpha)\%$

  · Given a single sample, $Y_1 = y_1, \ldots, Y_n = y_n$, suppose one has a procedure generating an intervals for $\theta$ with the confidence $100(1-\alpha)\%$.

· Then an interval based on a single sample, $y_1, \ldots, y_n$, either contains the true value of $\theta$ or it does not.

· If the same procedure was used many times (in repeated sampling), each interval would either contain or fail to contain the true value of $\theta$. But, the *percentage* of all intervals enclosing the true value of $\theta$ would be very close to $100(1-\alpha)\%$.

· $100(1-\alpha)\%$ *confidence intervals*: interval estimators with the confidence $100(1-\alpha)\%$

- Goal : construct a confidence interval, $[\widehat{\theta}_L, \widehat{\theta}_U]$ that can generate narrow intervals having a high probability of enclosing $\theta$ (under repeated sampling).

- Two-sided confidence interval, $[\widehat{\theta}_L, \widehat{\theta}_U]$

  · $\widehat{\theta}_L$ and $\widehat{\theta}_U$ : random lower and upper end-point (i.e., confidence limit)

  · $1-\alpha = P(\widehat{\theta}_L \leq \theta \leq \widehat{\theta}_U)$, confidence coefficient

  · The probability that an interval based on a single sample of size $n$ would contain $\theta$ is zero or one. If the same procedure were implemented many times in repeated sampling, each individual interval would either contain or fail to contain $\theta$, but the fraction of the time that the constructed intervals will contain the true value of $\theta$ would be close to $100(1-\alpha)$.

  · prefer intervals with higher $(1-\alpha)$ if their lengths are the same.

  · prefer narrower confidence intervals with the same $(1-\alpha)$.

- Lower one-sided confidence interval, $[\widehat{\theta}_L, \infty)$

  · $P(\widehat{\theta}_L \leq \theta) = 1-\alpha$

- Upper one-sided confidence interval, $(-\infty, \widehat{\theta}_U]$

  · $P(\theta \leq \widehat{\theta}_U) = 1-\alpha$

- How to find confidence intervals?

  · Use pivotal method : need to find a pivotal quantity having two characteristics

  i) it is a function of $Y_1, \ldots, Y_n$ and unknown $\theta$ where $\theta$ is the only unknown quantity.

  ii) its probability distribution does not depend on $\theta$

  · Logic of the pivotal method: for a r.v. $Y$, suppose that the probability distribution of the pivotal quantity is known. Then, $P(a \leq Y \leq b) = P\left(c(a+d) \leq a(Y+d) \leq c(b+d)\right)$

(Example 8.4)

(Example 8.5)

(Exercise 8.46)

## 8.6 Large-sample confidence intervals

- Pivotal method to develop confidence intervals for $\theta$ when sample size is large

  · Approximate probability distribution of $\frac{\widehat{\theta}-\theta}{\sigma_{\widehat{\theta}}}$ is a standard normal distribution, $N(0,1)$ as long as sample size is large and $E(\widehat{\theta}) = \theta$ (note that this holds for four unbiased estimators in Table 8.1).

  · Then a $100(1-\alpha)\%$ two-sided confidence interval for $\theta$ is

  $$[\widehat{\theta}_L, \widehat{\theta}_U] = [\widehat{\theta} - z_{\alpha/2}\sigma_{\widehat{\theta}}, \widehat{\theta} + z_{\alpha/2}\sigma_{\widehat{\theta}}]$$

  where $z_{\alpha/2}$ is a value satisfying $P(Z \geq z_{\alpha/2}) = \alpha/2$ and $Z \sim N(0,1)$

  Why? Since $\frac{\widehat{\theta}-\theta}{\sigma_{\widehat{\theta}}} \sim N(0,1)$,

  $$\begin{aligned} 1-\alpha &= P(-z_{\alpha/2} \leq \frac{\widehat{\theta}-\theta}{\sigma_{\widehat{\theta}}} \leq z_{\alpha/2}) \\ &= P(-z_{\alpha/2}\sigma_{\widehat{\theta}} \leq \widehat{\theta} - \theta \leq z_{\alpha/2}\sigma_{\widehat{\theta}}) \\ &= P(\widehat{\theta} - z_{\alpha/2}\sigma_{\widehat{\theta}} \leq \theta \leq \widehat{\theta} + z_{\alpha/2}\sigma_{\widehat{\theta}}) \end{aligned}$$

  How about a $100(1-\alpha)\%$ one-sided interval for $\theta$?

· Unknown $\sigma^2_{\widehat{\theta}}$ might be replaced with estimated value, $\widehat{\sigma}^2_{\widehat{\theta}}$ as long as $n$ is large. (note that the calculated confidence interval will have *approximately* the stated confidence coefficient).

For $\theta = \mu$, $\widehat{\theta} = \bar{Y}$ and $\sigma^2_{\widehat{\theta}} = \sigma^2/n$. Use sample variance, $s^2$ for unknown $\sigma^2$.
For $\theta = p$, $\widehat{\theta} = \widehat{p}$ and $\sigma^2_{\widehat{\theta}} = \sqrt{pq/n}$. Use $\widehat{p}$ for unknown $p$.

(Example 8.7)

(Example 8.8)

(Exercise 8.56) In a Gallup Poll of $n = 800$ randomly chosen adults, 45% indicated that movies were getting better whereas 43% indicated that movies were getting worse.
**(a)** Find a 98% confidence interval for $p$, the overall proportion of adults who say that movies are getting better?

## 8.7 Selecting the sample size

- Method of choosing $n$ using the large-sample confidence intervals procedure: $\frac{\widehat{\theta}-\theta}{\sigma_{\widehat{\theta}}} \overset{\cdot}{\sim} N(0,1)$,

  1) The following information should be given by the experimenter : a desired bound, $B$ on the error of estimation, and an associated confidence level, $1-\alpha$.

  2) $1-\alpha = P(\mid \widehat{\theta}-\theta \mid \leq z_{\alpha/2}\sigma_{\widehat{\theta}})$ means $\alpha = P(\mid \widehat{\theta}-\theta \mid > z_{\alpha/2}\sigma_{\widehat{\theta}})$

  3) Calculate (approximate) $n$ by equating $z_{\alpha/2}\sigma_{\widehat{\theta}} = B$ satisfying $P(Z > z_{\alpha/2}) = \alpha/2$ where $Z \sim N(0,1)$. For unknown $\sigma_{\widehat{\theta}}$ use one of the following methods

  · replace unknown $\sigma_{\widehat{\theta}}$ with its estimate $\widehat{\sigma}_{\widehat{\theta}}$

  · $\sigma_{\widehat{\theta}}$ =range/4 where the *range* of a set of measurement is the difference between the largest and smallest values (this works if the distribution of measurements is approximately normal).

(Example) suppose that one wants to estimate the average daily yield $\mu$ of a chemical. If one wishes the error of estimation to be less than 5 tons with probability .95, how large $n$ should be? Assume also that the range of the daily yields is known to be approximately 84 tons.

(Example 8.9) The reaction of an individual to a stimulus in a psychological experiment may take one of two forms, A and B, If an experiment wishes to estimate the probability $p$ that a person will react in manner A, how many people must be included in the experiment? Assume that the experimenter will be satisfied if the error of estimation is less than .04 with probability equal to .90. Assume also that he expects $p$ to be close to .6.

(Example 8.10)

## 8.8 Small-sample confidence intervals for $\mu$, and $\mu_1 - \mu_2$

[Case 1] Given $n$ random samples, $Y_1, \ldots, Y_n \sim N(\mu, \sigma^2)$, unknown $\sigma^2$, and small $n$, how to construct a confidence interval for $\mu$?.

[Case 2] Given $n_1$ random samples, $Y_{11}, \ldots, Y_{1n_1} \sim N(\mu_1, \sigma^2)$, $n_2$ random samples, $Y_{21}, \ldots, Y_{2n_2} \sim N(\mu_2, \sigma^2)$, unknown $\sigma^2$, and small $n_1$ and $n_2$, how to obtain a confidence interval for $\mu_1 - \mu_2$? (assume that they are independent samples).

· The large sample procedure in Chapter 8.8 might not be suitable.

· Use a $t$-distribution with parameter $\nu$(called degree of freedom), $T = \frac{Z}{\sqrt{W/\nu}} \sim t(\nu - 1)$ where $Z \sim N(0, 1)$, $W \sim \chi^2(\nu)$, and $Z$ and $W$ are independnt (see Def 7.2).

· $T$ is a pivotal quantity and $P(-t_{\alpha/2} \leq T \leq t_{\alpha/2}) = 1 - \alpha$. (see Table 5, T-table)

[Case 1] Given $n$ random samples, $Y_1, \ldots, Y_n \sim N(\mu, \sigma^2)$, unknown $\sigma^2$, and small $n$,

· $\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, and $\bar{Y}$ and $S^2$ are independent(Thm 7.1, 7.3)

· $T = \frac{\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\bar{Y}-\mu}{S/\sqrt{n}} \sim t(n-1)$ where $S = \sqrt{S^2}$

· The $100(1-\alpha)\%$ two-sided confidence interval for $\mu$ is

$$[\bar{Y} - t_{\alpha/2}(S/\sqrt{n}), \bar{Y} + t_{\alpha/2}(S/\sqrt{n})].$$

Why? $1-\alpha = P(-t_{\alpha/2} \le T \le t_{\alpha/2}) = P(\bar{Y} - t_{\alpha/2}(S/\sqrt{n}) \le \mu \le \bar{Y} + t_{\alpha/2}(S/\sqrt{n}))$

(Example 8.11) suppose a manufacturer of gunpowder has developed a new powder, which was tested in eight shells, and measured the muzzle velocity, $Y_i$, $i = 1, \ldots, 8$. Assume that $Y_i$ were normally distributed, its variance $\sigma^2$ was unknown, and their sample mean and sample variance were $\bar{y} = 2959$ and $s^2 = 39.1^2$. Find a 95% confidence interval for the true average velocity $\mu$ for shells of this type.

[Case 2] Given two sets of random samples, $Y_{11}, \ldots, Y_{1n_1} \sim N(\mu_1, \sigma^2)$, $Y_{21}, \ldots, Y_{2n_2} \sim N(\mu_2, \sigma^2)$, unknown common $\sigma^2$, and small $n_1$ and $n_2$,

- $T = \dfrac{\frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2/n_1 + \sigma^2/n_2}}}{\sqrt{\frac{(n_1+n_2-2)S_p^2}{\sigma^2(n_1+n_2-2)}}} = \dfrac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}} \sim t(n_1+n_2-2)$ where $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$,

- The $100(1-\alpha)\%$ two-sided confidence interval for $\mu$ is

$$\left[(\bar{Y}_1 - \bar{Y}_2) - t_{\alpha/2} S_p \sqrt{1/n_1 + 1/n_2},\right.$$
$$\left.(\bar{Y}_1 - \bar{Y}_2) + t_{\alpha/2} S_p \sqrt{1/n_1 + 1/n_2}\right]$$

(Example 8.12)

(Exercise 8.83)

## 8.9 Confidence intervals for $\sigma^2$

- Population variance $\sigma^2$ : amount of variability in the population
- An unbiased estimator for unknown $\sigma^2$ : $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ (example 8.1)
- Confidence interval for $\sigma^2$ using the pivotal method

  · Suppose a random sample, $Y_1, \ldots, Y_n \sim N(\mu, \sigma^2)$ where $\mu$ and $\sigma^2$ are unknown.

  · Then a $100(1-\alpha)\%$ two-sided interval is $[\sigma_L^2, \sigma_U^2] = \left[\frac{(n-1)S^2}{\chi^2_{(\alpha/2)}}, \frac{(n-1)S^2}{\chi^2_{1-(\alpha/2)}}\right]$ (see Table 6)

  Why? since $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$(Thm 7.3),

$$1 - \alpha = P\left[\chi^2_{1-(\alpha/2)} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{(\alpha/2)}\right]$$
$$= P\left[\frac{(n-1)S^2}{\chi^2_{(\alpha/2)}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-(\alpha/2)}}\right]$$

- How about a $100(1-\alpha)\%$ one-sided interval for $\sigma^2$?

(Example 8.13) An experimenter wanted to check the variability of measurements obtained by using equipment designed to measure the volume of an audio source. Three independent measurements recorded by this equipment for the same sound were 4.1, 5.2 and 10.2. Estimate $\sigma^2$ with confidence coefficient .90.