ST 301 (AKI) LECTURE #23

1 Review: Ch9.5 Inferences about the $\sigma$

(p385)

Eg) Let $\bar{x} = 0.7$, $s = 0.4$, $n=10$.

Q1) Do $H_0 : \sigma^2 = 0.2$, $H_1 : \sigma^2 < 0.2$ with $\alpha = 0.1$

Q2) Find 95% C.I. for $\sigma$

Hint:

$qchisq(0.1, df = 9) = 4.17$,

$qchisq(0.9, df = 9) = 14.68$
2 Ch10 Comparing Two Treatment (p398)

2.1 Terminology (p398-p399)

- **Treatment**: the things that are being compared.
  eg) Water and coke.

- **Experimental Unit**: the basic unit or person that is exposed to one treatment or another.
  eg) Person.

- **Response**: the characteristic that is recorded after the application of a treatment to a subject
  eg) Sleeping time.

- **Experimental Design**: the manner in which subjects are chosen and assigned to treatments.
2.2 Large samples from two populations (p403-)

- $X_1, X_2, \ldots, X_{n_1}$ is a random sample of size $n_1$ from population 1 with $E[X_1] = \mu_1$, $Var[X_1] = \sigma_1^2$.

- $Y_1, Y_2, \ldots, Y_{n_2}$ is a random sample of size $n_2$ from population 2 with $E[Y_1] = \mu_2$, $Var[Y_1] = \sigma_2^2$.

- The samples are independent. In other words, the response measurements under one treatment are unrelated to the response measurements under the other treatment.

- $n_1 > 30$, $n_2 > 30$. 
The parameter of interest: $\mu_1 - \mu_2$. Note that if those two populations have the same mean, then

$$\mu_1 - \mu_2 = 0. \quad (1)$$

If not, we have

$$\mu_1 - \mu_2 = \delta_0. \quad (2)$$

(for example $\mu_1 - \mu_2 = 3.$) Here we have two parameters. But we can think

$$\mu \equiv \mu_1 - \mu_2, \quad (3)$$

and we are doing hypothesis test for one parameter $\mu$ as we learned in ch8, ch9. Then we can use a similar method for hypothesis test and C.I.

Estimator of $\mu_1 - \mu_2$: $\bar{X} - \bar{Y}$

Mean:

$$E[\bar{X} - \bar{Y}] = \mu_1 - \mu_2. \quad (4)$$

Variance:

$$V ar[\bar{X} - \bar{Y}] = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \quad (5)$$

Standard Error:

$$S.E. [\bar{X} - \bar{Y}] = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (6)$$
100(1 - \alpha)\% CI of \mu_1 - \mu_2 is

\left( \overline{X} - \overline{Y} - z_{\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}, \overline{X} - \overline{Y} + z_{\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \right),

(7)

where

\begin{align*}
S_1^2 &= \frac{\sum_{i=1}^{n_1} (X_i - \overline{X})^2}{n_1 - 1} \quad (8) \\
S_2^2 &= \frac{\sum_{i=1}^{n_2} (Y_i - \overline{Y})^2}{n_2 - 1} \quad (9)
\end{align*}
Eg1) Independent random samples from two populations have provided the summary statistics as follows:

Sample 1: \( n_1 = 52, \bar{x} = 73, s_1^2 = 151 \)
Sample 2: \( n_2 = 44, \bar{y} = 66, s_1^2 = 142 \)

Q1) Obtain a point estimate of \( \mu_1 - \mu_2 \)

Q2) Obtain the estimated standard error.

Q3) Construct 95% CI for \( \mu_1 - \mu_2 \)
Q4) Do hypothesis test with $\alpha = 0.05$, and find p-value.
$H_0 : \mu_1 - \mu_2 = 6,$
$H_1 : \mu_1 - \mu_2 > 6.$ (C1 test)

Q5) Do hypothesis test with $\alpha = 0.05$, and find p-value.
$H_0 : \mu_1 - \mu_2 = 6,$
$H_1 : \mu_1 - \mu_2 < 6$ (C2 test)
Q6) Do hypothesis test with $\alpha = 0.05$, and find p-value.
$H_0 : \mu_1 - \mu_2 = 6$,
$H_1 : \mu_1 - \mu_2 \neq 6$ (C3 test)
2.3 Small samples from two NORMAL populations and equal variance

- $X_1, X_2, \cdots, X_{n_1}$ is a random sample of size $n_1$ from population 1 with $N(\mu_1, \sigma_1^2)$.

- $Y_1, Y_2, \cdots, Y_{n_2}$ is a random sample of size $n_2$ from population 2 with $N(\mu_2, \sigma_2^2)$.

- The samples are independent. In other words, the response measurements under one treatment are unrelated to the response measurements under the other treatment.

- Small sample case.

- $1/2 \leq s_1/s_2 \leq 2$ (i.e. we can assume $\sigma_1 = \sigma_2$).

Note: (p412) said ”Assume $\sigma_1 = \sigma_2$ ”. But in reality, we do not know the population variance, hence we need to check $1/2 \leq s_1/s_2 \leq 2$. If it does not hold, we need another method to do hypothesis test, as stated in the next section.
Estimator of $\mu_1 - \mu_2$: $\overline{X} - \overline{Y}$ (same as the previous section)

100$(1 - \alpha)$% CI of $\mu_1 - \mu_2$ is

$$\overline{X} - \overline{Y} \pm t_{df=n_1+n_2-2,\alpha/2} \cdot S_{\text{Pooled}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$  \hspace{1cm} (10)

where

$$S_{\text{Pooled}}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$  \hspace{1cm} (11)
Hypothesis test with $\alpha$ level (C3):

$H_0 : \mu_1 - \mu_2 = \delta_0$

$H_1 : \mu_1 - \mu_2 \neq \delta_0$

Step1) Get

$$T_{obs} = \frac{\overline{X} - \overline{Y} - \delta_0}{S_{Pooled} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$  \hspace{1cm} (12)$$

If $T_{obs}$ is in rejection region, ”Reject $H_0.”$ In other word, if $|T_{obs}| > t_{df=n_1+n_2-2, \alpha/2}$, then ”Reject $H_0.”$ 

Suppose you do $H_1 : \mu_1 - \mu_2 > \delta_0$ (C1) test.
If $T_{obs} > t_{df=n_1+n_2-2, \alpha}$, then ”Reject $H_0.”$ 

Suppose you do $H_1 : \mu_1 - \mu_2 < \delta_0$ (C2) test.
If $T_{obs} < -t_{df=n_1+n_2-2, \alpha}$, then ”Reject $H_0.”$ 

• p-value for C1, C2, C3 (Figures:)