STATISTICAL INFERENCES (2cr)
Chapter 9 One- and Two-Sample Estimation Problems

Zhong Guan

Math, IUSB

Outline

Outline

Assumptions and Point Estimator

- Assume that $X_1, \ldots, X_n$ is a random sample from normal population with mean μ and variance σ^2.
- Both μ and σ^2 are unknown.
- The point estimator is the sample variance $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$. This is an unbiased and consistent estimator.

Assumptions and Point Estimator

- Assume that $X_1, \ldots, X_n$ is a random sample from normal population with mean μ and variance σ^2.
- Both μ and σ^2 are unknown.
- The point estimator is the sample variance $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$. This is an unbiased and consistent estimator.

Assumptions and Point Estimator

- Assume that $X_1, \ldots, X_n$ is a random sample from normal population with mean μ and variance σ^2.
- Both μ and σ^2 are unknown.
- The point estimator is the sample variance $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$. This is an unbiased and consistent estimator.

Assumptions and Point Estimator

- Assume that $X_1, \ldots, X_n$ is a random sample from normal population with mean μ and variance σ^2.
- Both μ and σ^2 are unknown.
- The point estimator is the sample variance $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$. This is an unbiased and consistent estimator.

Confidence Interval

- $X^2 = (n-1)S^2/\sigma^2$ has a chi-squared distribution with $n-1$ degrees of freedom.
- There exist a and b such that

$$1-\alpha = P(a \leqslant X^2 \leqslant b) = P(a \leqslant (n-1)S^2/\sigma^2 \leqslant b)$$
$$= P((n-1)S^2/b \leqslant \sigma^2 \leqslant (n-1)S^2/a)$$

- For convenience we choose $a = \chi^2_{1-\alpha/2}$ and $b = \chi^2_{\alpha/2}$. Thus we have a $100(1-\alpha)\%$ confidence interval for σ^2:

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

Confidence Interval

- $X^2 = (n-1)S^2/\sigma^2$ has a chi-squared distribution with $n-1$ degrees of freedom.
- There exist a and b such that

$$1-\alpha = P(a \leqslant X^2 \leqslant b) = P(a \leqslant (n-1)S^2/\sigma^2 \leqslant b)$$
$$= P((n-1)S^2/b \leqslant \sigma^2 \leqslant (n-1)S^2/a)$$

- For convenience we choose $a = \chi^2_{1-\alpha/2}$ and $b = \chi^2_{\alpha/2}$. Thus we have a $100(1-\alpha)\%$ confidence interval for σ^2:

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

Confidence Interval

- $X^2 = (n-1)S^2/\sigma^2$ has a chi-squared distribution with $n-1$ degrees of freedom.
- There exist a and b such that

$$\begin{aligned} 1-\alpha &= P(a \leqslant X^2 \leqslant b) = P(a \leqslant (n-1)S^2/\sigma^2 \leqslant b) \\ &= P((n-1)S^2/b \leqslant \sigma^2 \leqslant (n-1)S^2/a) \end{aligned}$$

- For convenience we choose $a = \chi^2_{1-\alpha/2}$ and $b = \chi^2_{\alpha/2}$. Thus we have a $100(1-\alpha)\%$ confidence interval for σ^2:

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

Find Critical Value χ^2_α

- Using Table A.5.
- Using TI 8x:
  (1) graph cdf curve $y = \chi^2 cdf(0, X, df)$ and horizontal line $y = 1 - \alpha$;
  (2) the x-coordinate of the intersection is χ^2_α.

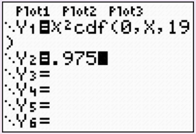
Find Critical Value χ^2_α

- Using Table A.5.
- Using TI 8x:
  (1) graph cdf curve $y = \chi^2 cdf(0, X, df)$ and horizontal line $y = 1 - \alpha$;
  (2) the x-coordinate of the intersection is χ^2_α.

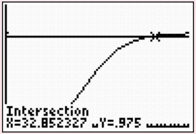
Use TI-83/84 to Find Critical χ^2 Value

For example, $n = 20$, df $=n - 1$=19, to find $\chi^2_{.025}$:

- Enter equation for χ^2 cdf : $\boxed{\text{Y=}}$ -> Y1= $\chi^2\text{cdf}(0, X, 19)$
- Enter equation: Y2=1-.025=0.975.



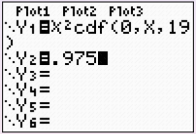
- Press $\boxed{\text{GRAPH}}$ and adjust $\boxed{\text{WINDOW}}$ if needed.
- Press $\boxed{\text{2ND}}$ and $\boxed{\text{CALC}}$ to find intersection of the two graphs. We can obtain $\chi^2_{.025} = 32.85$.



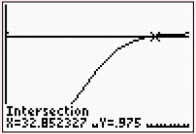
Use TI-83/84 to Find Critical χ^2 Value

For example, $n = 20$, df $= n - 1 = 19$, to find $\chi^2_{.025}$:

- Enter equation for χ^2 cdf : $\boxed{\text{Y=}}$ -> Y1= χ^2cdf(0, X, 19)
- Enter equation: Y2=1-.025=0.975.



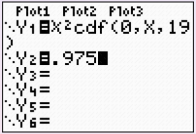
- Press $\boxed{\text{GRAPH}}$ and adjust $\boxed{\text{WINDOW}}$ if needed.
- Press $\boxed{\text{2ND}}$ and $\boxed{\text{CALC}}$ to find intersection of the two graphs. We can obtain $\chi^2_{.025} = 32.85$.



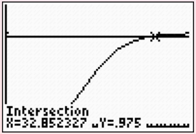
Use TI-83/84 to Find Critical χ^2 Value

For example, $n = 20$, df $= n - 1 = 19$, to find $\chi^2_{.025}$:

- Enter equation for χ^2 cdf : $\boxed{\text{Y=}}$ -> Y1= $\chi^2\text{cdf}$(0, X, 19)
- Enter equation: Y2=1-.025=0.975.



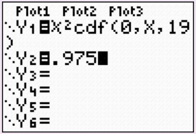
- Press $\boxed{\text{GRAPH}}$ and adjust $\boxed{\text{WINDOW}}$ if needed.
- Press $\boxed{\text{2ND}}$ and $\boxed{\text{CALC}}$ to find intersection of the two graphs. We can obtain $\chi^2_{.025} = 32.85$.



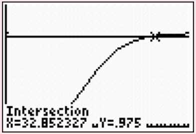
Use TI-83/84 to Find Critical χ^2 Value

For example, $n = 20$, df $= n - 1 = 19$, to find $\chi^2_{.025}$:

- Enter equation for χ^2 cdf : $\boxed{\text{Y=}}$ -> Y1= $\chi^2\mathrm{cdf}(0, X, 19)$
- Enter equation: Y2=1-.025=0.975.



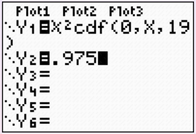
- Press $\boxed{\text{GRAPH}}$ and adjust $\boxed{\text{WINDOW}}$ if needed.
- Press $\boxed{\text{2ND}}$ and $\boxed{\text{CALC}}$ to find intersection of the two graphs. We can obtain $\chi^2_{.025} = 32.85$.



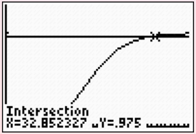
Use TI-83/84 to Find Critical χ^2 Value

For example, $n = 20$, df $= n - 1 = 19$, to find $\chi^2_{.025}$:

- Enter equation for χ^2 cdf : $\boxed{\text{Y=}}$ -> Y1= $\chi^2\mathrm{cdf}(0, X, 19)$
- Enter equation: Y2=1-.025=0.975.



- Press $\boxed{\text{GRAPH}}$ and adjust $\boxed{\text{WINDOW}}$ if needed.
- Press $\boxed{\text{2ND}}$ and $\boxed{\text{CALC}}$ to find intersection of the two graphs. We can obtain $\chi^2_{.025} = 32.85$.



Challenging Problems

- Find a and b so that
$$1-\alpha = P((n-1)S^2/b \leqslant \sigma^2 \leqslant (n-1)S^2/a)$$
and $\frac{1}{a}-\frac{1}{b}$ is minimized.
- If μ is known, how do you estimate σ^2 using point estimator and confidence interval?

Assumptions and Point Estimator

- Assume that $X_{11}, \ldots, X_{1n_1}$ is a random sample from normal population with unknown mean μ_1 and variance σ_1^2, and
- $X_{21}, \ldots, X_{2n_2}$ is a random sample from normal population with unknown mean μ_2 and variance σ_2^2, and that
- the two samples are independent.
- The sample variances and sample means are

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}, \quad i = 1, 2.$$

- The point estimator of σ_1^2/σ_2^2 is S_1^2/S_2^2. This is a consistent but biased estimator.

Assumptions and Point Estimator

- Assume that $X_{11}, \ldots, X_{1n_1}$ is a random sample from normal population with unknown mean μ_1 and variance σ_1^2, and
- $X_{21}, \ldots, X_{2n_2}$ is a random sample from normal population with unknown mean μ_2 and variance σ_2^2, and that
- the two samples are independent.
- The sample variances and sample means are

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}, \quad i = 1, 2.$$

- The point estimator of σ_1^2/σ_2^2 is S_1^2/S_2^2. This is a consistent but biased estimator.

Assumptions and Point Estimator

- Assume that $X_{11}, \ldots, X_{1n_1}$ is a random sample from normal population with unknown mean μ_1 and variance σ_1^2, and
- $X_{21}, \ldots, X_{2n_2}$ is a random sample from normal population with unknown mean μ_2 and variance σ_2^2, and that
- the two samples are independent.
- The sample variances and sample means are

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}, \quad i = 1, 2.$$

- The point estimator of σ_1^2/σ_2^2 is S_1^2/S_2^2, This is a consistent but biased estimator.

Assumptions and Point Estimator

- Assume that $X_{11}, \ldots, X_{1n_1}$ is a random sample from normal population with unknown mean μ_1 and variance σ_1^2, and
- $X_{21}, \ldots, X_{2n_2}$ is a random sample from normal population with unknown mean μ_2 and variance σ_2^2, and that
- the two samples are independent.
- The sample variances and sample means are

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}, \quad i = 1, 2.$$

- The point estimator of σ_1^2/σ_2^2 is S_1^2/S_2^2, This is a consistent but biased estimator.

Assumptions and Point Estimator

- Assume that $X_{11}, \ldots, X_{1n_1}$ is a random sample from normal population with unknown mean μ_1 and variance σ_1^2, and
- $X_{21}, \ldots, X_{2n_2}$ is a random sample from normal population with unknown mean μ_2 and variance σ_2^2, and that
- the two samples are independent.
- The sample variances and sample means are

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}, \quad i = 1, 2.$$

- The point estimator of σ_1^2/σ_2^2 is S_1^2/S_2^2, This is a consistent but biased estimator.

Assumptions and Point Estimator

- Assume that $X_{11}, \ldots, X_{1n_1}$ is a random sample from normal population with unknown mean μ_1 and variance σ_1^2, and
- $X_{21}, \ldots, X_{2n_2}$ is a random sample from normal population with unknown mean μ_2 and variance σ_2^2, and that
- the two samples are independent.
- The sample variances and sample means are

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}, \quad i = 1, 2.$$

- The point estimator of σ_1^2/σ_2^2 is S_1^2/S_2^2, This is a consistent but biased estimator.

Confidence Interval

- $X_1^2 = (n_1 - 1)S_1^2/\sigma_1^2$ and $X_2^2 = (n_2 - 1)S_2^2/\sigma_2^2$ are independent and have chi-squared distributions with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom, respectively.
- So $F = \frac{X_1^2/v_1}{X_2^2/v_2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$ has an F distribution with numerator and denominator degrees of freedom $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$, respectively.

Confidence Interval

- $X_1^2 = (n_1 - 1)S_1^2/\sigma_1^2$ and $X_2^2 = (n_2 - 1)S_2^2/\sigma_2^2$ are independent and have chi-squared distributions with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom, respectively.
- So $F = \frac{X_1^2/v_1}{X_2^2/v_2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$ has an F distribution with numerator and denominator degrees of freedom $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$, respectively.

Confidence Interval

- $X_1^2 = (n_1 - 1)S_1^2/\sigma_1^2$ and $X_2^2 = (n_2 - 1)S_2^2/\sigma_2^2$ are independent and have chi-squared distributions with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom, respectively.
- So $F = \frac{X_1^2/v_1}{X_2^2/v_2} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$ has an F distribution with numerator and denominator degrees of freedom $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$, respectively.

Confidence Interval(Cont.)

- There exist a and b such that

$$1-\alpha = P(a \leqslant F \leqslant b) = P\left(a \leqslant \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \leqslant b\right)$$

$$= P\left(\frac{S_1^2}{bS_2^2} \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{aS_2^2}\right)$$

- For convenience we choose $a = F_{1-\alpha/2}$ and $b = F_{\alpha/2}$. Thus we have a $100(1-\alpha)\%$ confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$:

$$\left[\frac{S_1^2}{F_{\alpha/2}S_2^2}, \frac{S_1^2}{F_{1-\alpha/2}S_2^2}\right]$$

Confidence Interval(Cont.)

- There exist a and b such that

$$1-\alpha = P(a \leqslant F \leqslant b) = P(a \leqslant \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \leqslant b)$$

$$= P(\frac{S_1^2}{bS_2^2} \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{aS_2^2})$$

- For convenience we choose $a = F_{1-\alpha/2}$ and $b = F_{\alpha/2}$. Thus we have a $100(1-\alpha)\%$ confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$:

$$\left[\frac{S_1^2}{F_{\alpha/2}S_2^2}, \frac{S_1^2}{F_{1-\alpha/2}S_2^2}\right]$$

Find Critical Value F_α

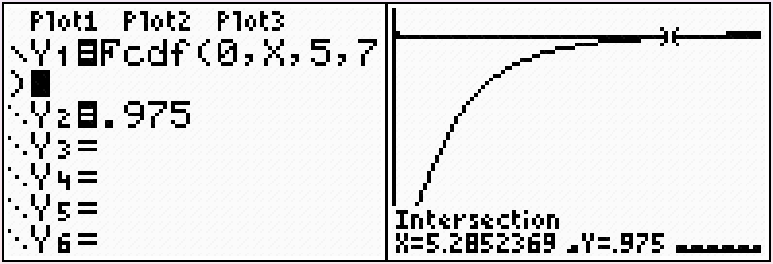
- Using Table A.6.
- Using TI 8x:
  (1) graph Fcdf $Y = Fcdf(0, X, v_1, v_2)$ and horizontal line $Y = 1 - \alpha$;
  (2) the x-coordinate of the intersection is F_α.

Find Critical Value F_α

- Using Table A.6.
- Using TI 8x:
  (1) graph Fcdf $Y = Fcdf(0, X, v_1, v_2)$ and horizontal line $Y = 1 - \alpha$;
  (2) the x-coordinate of the intersection is F_α.

The F Critical Value using Technology

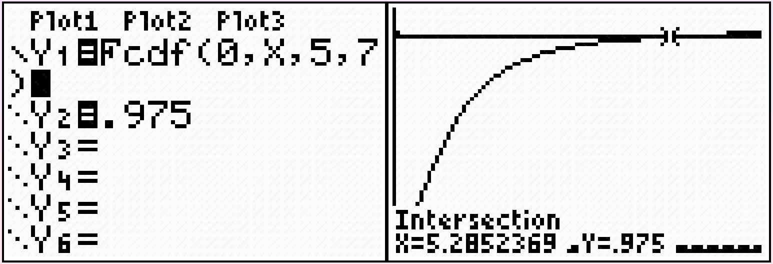
- In Excel: $F_{.025}$=F.INV.RT(0.025,5,7)=5.285236852, or =F.INV(1-0.025,5,7)=5.285236852.
- In TI-8x: Graph Y1=Fcdf(0, X, 5,7) and Y2=1-.025=0.975, the x-coordinate of the intersection is $F_{.025}$=5.2852369.



- In R: $F_{.975}$= "qf(.975,5,7)=5.285237".

The F Critical Value using Technology

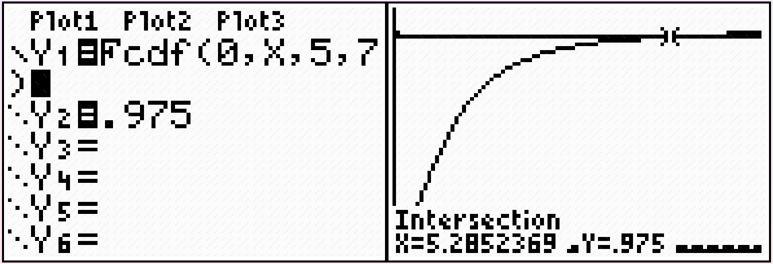
- In Excel: $F_{.025}$=F.INV.RT(0.025,5,7)=5.285236852, or =F.INV(1-0.025,5,7)=5.285236852.
- In TI-8x: Graph Y1=Fcdf(0, X, 5,7) and Y2=1-.025=0.975, the x-coordinate of the intersection is $F_{.025}$=5.2852369.



- In R: $F_{.975}$= "qf(.975,5,7)=5.285237".

The F Critical Value using Technology

- In Excel: $F_{.025}$=F.INV.RT(0.025,5,7)=5.285236852, or =F.INV(1-0.025,5,7)=5.285236852.
- In TI-8x: Graph Y1=Fcdf(0, X, 5,7) and Y2=1-.025=0.975, the x-coordinate of the intersection is $F_{.025}$=5.2852369.



- In R: $F_{.975}$= “qf(.975,5,7)=5.285237”.

Challenging Problems

- Find a and b so that

$$1-\alpha = P\left(\frac{S_1^2}{bS_2^2} \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{aS_2^2}\right)$$

  and $\frac{1}{a} - \frac{1}{b}$ is minimized.
- If μ_1 and μ_2 are known, how do you estimate σ_1^2/σ_2^2 using point estimator and confidence interval?