Simultaneous Inference: An Overview

Topics to be covered:

- Joint estimation of β_0 and β_1.
- Simultaneous estimation of mean responses.
- Simultaneous prediction intervals.

Statement Confidence Coefficient

- We learned how to construct individual CIs for β_0 and β_1.
- Statement confidence coefficient: Indicates the proportion of correct estimates that are obtained when repeated samples are selected and the specified CI is calculated for each sample.
- However, individual CIs do not provide 95% confidence that the conclusions for **both** β_0 and β_1 are correct.
- For example, if inferences are independent, then

$$P(\text{Both CIs are correct}) = 0.95^2 = 0.9025.$$

- For 10 independent CIs,

$P(\text{At least one is wrong}) =$

Statement Confidence Coefficient

- We learned how to construct individual CIs for β_0 and β_1.
- Statement confidence coefficient: Indicates the proportion of correct estimates that are obtained when repeated samples are selected and the specified CI is calculated for each sample.
- However, individual CIs do not provide 95% confidence that the conclusions for **both** β_0 and β_1 are correct.
- For example, if inferences are independent, then

$$P(\text{Both CIs are correct}) = 0.95^2 = 0.9025.$$

- For 10 independent CIs,

$P(\text{At least one is wrong}) =$

Family Confidence Coefficient

- Family confidence coefficient: Indicates the proportion of families of estimates that are entirely correct when repeated samples are selected and the specified CIs for the entire family are calculated for each sample.
- For example, with a family confidence coefficient of 0.95 for β_0 and β_1, we can say that

  "In repeated samples, for 5% of the samples, either the CI for β_0, or the CI for β_1, or both CIs would not cover the true values.
- Are simultaneous intervals generally wider or narrower than individual CIs?

Bonferroni Idea

- Main idea: Each statement confidence coefficient is adjusted to be higher than $1-\alpha$, so the family confidence coefficient is at $1-\alpha$.
- Consider the individual $(1-\alpha)$ CI for β_0 and $(1-\alpha)$ CI for β_1.
- Let $A_1 =$ the event that CI for β_0 does not cover β_0.
- Let $A_2 =$ the event that CI for β_1 does not cover β_1.
- Thus $P(A_1) = \alpha$ and $P(A_2) = \alpha$.
- Bonferroni inequality (shown in the next slide):

$$P(\text{both cover}) \geq 1 - 2\alpha.$$

- Thus if each CI is constructed with $(1-\alpha/2)100\%$ confidence, then the joint confidence level will be at least $(1-\alpha)100\%$.

Bonferroni Idea

- Main idea: Each statement confidence coefficient is adjusted to be higher than $1-\alpha$, so the family confidence coefficient is at $1-\alpha$.
- Consider the individual $(1-\alpha)$ CI for β_0 and $(1-\alpha)$ CI for β_1.
- Let $A_1 =$ the event that CI for β_0 does not cover β_0.
- Let $A_2 =$ the event that CI for β_1 does not cover β_1.
- Thus $P(A_1) = \alpha$ and $P(A_2) = \alpha$.
- Bonferroni inequality (shown in the next slide):

$$P(\text{both cover}) \geq 1-2\alpha.$$

- Thus if each CI is constructed with $(1-\alpha/2)100\%$ confidence, then the joint confidence level will be at least $(1-\alpha)100\%$.

Bonferroni Justification

- Show the Bonferroni inequality for 2 CIs

Bonferroni Joint Confidence Intervals

- Bonferroni $(1-\alpha)$ joint CIs for β_0 and β_1 are

$$\begin{aligned}\hat{\beta}_0 &\pm t(1-\alpha/4; n-2)s\{\hat{\beta}_0\} \\ \hat{\beta}_1 &\pm t(1-\alpha/4; n-2)s\{\hat{\beta}_1\}\end{aligned}$$

- Interpretation: The family confidence coefficient is at least $(1-\alpha)$ that the procedure leads to the correct set of CIs.
- In the advertising example, recall that

$$\hat{\beta}_0 = 0.35, \quad s\{\hat{\beta}_0\} = 0.347, \quad \hat{\beta}_1 = 1.36, \quad s\{\hat{\beta}_1\} = 0.0886.$$

- At $\alpha = 0.10$, $t(1-\alpha/4; n-2) = t(0.975; 5) = 2.571$.
- Thus Bonferroni 90% joint CIs for β_0 and β_1 are

$$\begin{aligned}0.35 &\pm 2.571 \times 0.347 = 0.35 \pm 0.89 = (-0.54, 1.24) \\ 1.36 &\pm 2.571 \times 0.0886 = 1.36 \pm 0.23 = (1.13, 1.59)\end{aligned}$$

Bonferroni Joint Confidence Intervals

- Bonferroni $(1-\alpha)$ joint CIs for β_0 and β_1 are

$$\begin{aligned} \hat{\beta}_0 &\pm t(1-\alpha/4; n-2)s\{\hat{\beta}_0\} \\ \hat{\beta}_1 &\pm t(1-\alpha/4; n-2)s\{\hat{\beta}_1\} \end{aligned}$$

- Interpretation: The family confidence coefficient is at least $(1-\alpha)$ that the procedure leads to the correct set of CIs.
- In the advertising example, recall that

$$\hat{\beta}_0 = 0.35, \quad s\{\hat{\beta}_0\} = 0.347, \quad \hat{\beta}_1 = 1.36, \quad s\{\hat{\beta}_1\} = 0.0886.$$

- At $\alpha = 0.10$, $t(1-\alpha/4; n-2) = t(0.975; 5) = 2.571$.
- Thus Bonferroni 90% joint CIs for β_0 and β_1 are

$$\begin{aligned} 0.35 &\pm 2.571 \times 0.347 = 0.35 \pm 0.89 = (-0.54, 1.24) \\ 1.36 &\pm 2.571 \times 0.0886 = 1.36 \pm 0.23 = (1.13, 1.59) \end{aligned}$$

Bonferroni Joint Confidence Intervals

- In general, for $g \geq 2$ CIs with a family confidence coefficient of $1 - \alpha$, construct individual CIs with statement confidence coefficient $(1 - \alpha/g)$.
  - The $(1 - \alpha)$ level is a lower bound and the statements are correct more than $(1 - \alpha)100\%$ of the time.
  - Thus the Bonferroni procedure can be extremely conservative if g is large, producing overly wide intervals.
  - Therefore method is more suitable when g is quite small.
- The Bonferroni idea is easy to understand and can be applied in general to construct CIs for parameters, PIs for new observations, etc.

Simultaneous Estimation of Mean Responses

- We learned how to construct an individual CI for $E(Y_h)$ at a given level X_h and for a given statement confidence coefficient.
- Now consider simultaneous estimation of several $E(Y_h)$ at several levels of X_h for a given family confidence coefficient.
- Two methods:
  - Working-Hotelling procedure.
  - Bonferroni procedure.

Working-Hotelling Procedure

- Working-Hotelling $(1-\alpha)$ simultaneous CIs for g mean responses:

$$\hat{Y}_h \pm W s\{\hat{Y}_h\}$$

where

$$\begin{aligned} \hat{Y}_h &= \\ s\{\hat{Y}_h\} &= \\ W^2 &= \end{aligned}$$

- This has a $(1-\alpha)$ family confidence coefficient on the entire regression line.
- That is, the proportion of time the estimating procedure will yield a confidence band that covers the entire regression line is $(1-\alpha)$.

Properties of Working-Hotelling Procedure

- WH provides a $(1-\alpha)$ family confidence coefficient on the entire regression line.
  - That is, the proportion of time the estimating procedure will yield a confidence band that covers the entire regression line is $(1-\alpha)$.
- $W > t$ because confidence band must encompass the entire regression line, not merely $E(Y_h)$.
- Bands are parabolas.
- WH is not the only good choice, but it is straightforward.
- WH is similar to Scheffé procedure (later).

Example

- In the advertising example, estimate the expected sale at advertising expenditure levels 3, 5, and 7.
- Compute $W = \sqrt{2F(0.95; 2, 5)} = 3.402$ and

X_h	3	5	7
$\hat{Y}_h$	4.41	7.12	9.83
$s\{\hat{Y}_h\}$	0.164	0.207	0.349

- Working-Hotelling 95% simultaneous CIs for 3 mean responses are

$$\begin{aligned} 4.41 &\pm 3.402 \times 0.164 = 4.41 \pm 0.56 = (3.85, 4.97) \\ 7.12 &\pm 3.402 \times 0.207 = 7.12 \pm 0.70 = (6.42, 7.82) \\ 9.83 &\pm 3.402 \times 0.349 = 9.83 \pm 1.19 = (8.64, 11.02) \end{aligned}$$

Bonferroni Procedure

- Bonferroni $(1-\alpha)$ simultaneous CIs for g mean responses:

$$\hat{Y}_h \pm Bs\{\hat{Y}_h\}$$

where

$$\begin{aligned} \hat{Y}_h &= \hat{\beta}_0 + \hat{\beta}_1 X_h \\ s\{\hat{Y}_h\} &= \sqrt{MSE\left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]} \\ B &= \end{aligned}$$

Bonferroni Procedure

- In the advertising example, estimate the expected sale at advertising expenditure levels 3, 5, and 7.
- Compute $B = t(1 - 0.05/6; 5) = 3.534$ and

X_h	3	5	7
$\hat{Y}_h$	4.41	7.12	9.83
$s\{\hat{Y}_h\}$	0.164	0.207	0.349

- Bonferroni 95% simultaneous CIs for 3 mean responses are

$$\begin{aligned}
4.41 &\pm 3.534 \times 0.164 = 4.41 \pm 0.58 = (3.83, 4.99) \\
7.12 &\pm 3.534 \times 0.207 = 7.12 \pm 0.73 = (6.39, 7.85) \\
9.83 &\pm 3.534 \times 0.349 = 9.83 \pm 1.23 = (8.60, 11.06)
\end{aligned}$$

Simultaneous Estimation of Mean Responses: Remarks

- Both Working-Hotelling and Bonferroni procedures provide lower bounds to the actual family confidence coefficient.
- For larger g, the Working-Hotelling confidence limits are tighter than the Bonferroni limits.
- The Working-Hotelling procedure encompasses all possible levels of X.
- In this class, we focus on the Working-Hotelling procedure.

Simultaneous Prediction Intervals

- We learned how to construct an individual PI for $Y_{h(\text{new})}$ at a given level X_h and for a given statement confidence coefficient.
- Now consider simultaneous prediction of g new observations on Y at g different levels of X_h and for a given family confidence coefficient.
- Two methods:
  - Scheffé procedure.
  - Bonferroni procedure.
- Recall

$$\begin{aligned}\hat{Y}_h &= \hat{\beta}_0 + \hat{\beta}_1 X_h \\ s\{\text{pred}\} &= \sqrt{MSE\left[1 + \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right]}\end{aligned}$$

Simultaneous PIs for g Predictions

- Let the family confidence coefficient be $(1-\alpha)$.
- Scheffé procedure:

$$\hat{Y}_h \pm Ss\{\text{pred}\}$$

where

- Bonferroni procedure:

$$\hat{Y}_h \pm Bs\{\text{pred}\}$$

where

Simultaneous PIs for g Predictions

- In the advertising example, predict new observations of sale at advertising expenditure levels 3, 5, and 7.
- Compute

X_h	3	5	7
$\hat{Y}_h$	4.41	7.12	9.83
$s\{\text{pred}\}$	0.450	0.467	0.545

Simultaneous PIs for g Predictions

- With $S = \sqrt{3F(0.95;3,5)} = 4.028$, Scheffé 95% simultaneous PIs are

$$4.41 \pm 4.028 \times 0.450 = 4.41 \pm 1.81 = (2.60, 6.22)$$
$$7.12 \pm 4.028 \times 0.467 = 7.12 \pm 1.88 = (5.24, 9.00)$$
$$9.83 \pm 4.028 \times 0.545 = 9.83 \pm 2.20 = (7.63, 12.03)$$

- With $B = t(1 - 0.05/6; 5) = 3.534$, Bonferroni 95% simultaneous PIs are

$$4.41 \pm 3.534 \times 0.450 = 4.41 \pm 1.54 = (2.82, 6.00)$$
$$7.12 \pm 3.534 \times 0.467 = 7.12 \pm 1.65 = (5.47, 8.77)$$
$$9.83 \pm 3.534 \times 0.545 = 9.83 \pm 1.93 = (7.90, 11.76)$$

Simultaneous Prediction Intervals: Remarks

- For a given α, the simultaneous PIs for g new observations are wider than individual PIs.
- Of course, the prediction intervals are wider than the confidence intervals for mean responses
- For prediction, both B and S increase as g increases, unlike the simultaneous CIs for the mean responses where B increases with g but W does not
- Both B and S can be (very) conservative as g increases