Random Explanatory Variables[1]

STA302 Fall 2020

[1] See last slide for copyright information.

Overview

Preparation: Change of Variables Formula

$Y = g(X)$

Two ways of writing the same thing:

$$
\begin{aligned}
E(Y) &= \int y f_Y(y)\, dy \\
E(g(X)) &= \int g(x) f_X(x)\, dx
\end{aligned}
$$

Preparation: Indicator functions

Conditional expectation and the Law of Total Probability

$I_A(x)$ is the *indicator function* for the set A. It is defined by

$$I_A(x) = \begin{cases} 1 & \text{for } x \in A \\ 0 & \text{for } x \notin A \end{cases}$$

Also sometimes written $I(x \in A)$

$$\begin{aligned} E(I_A(X)) &= \sum_x I_A(x)p(x) = \sum_{x \in A} p(x), \text{ or} \\ & \int_{-\infty}^{\infty} I_A(x) f(x)\, dx = \int_A f(x)\, dx \\ &= P\{X \in A\} \end{aligned}$$

So the expected value of an indicator is a probability.

Applies to conditional probabilities too

$$
\begin{aligned}
E(I_A(X)|Y) &= \sum_x I_A(x)p(x|Y), \text{ or} \\
& \int_{-\infty}^{\infty} I_A(x)f(x|Y)\,dx \\
&= Pr\{X \in A|Y\}
\end{aligned}
$$

So the conditional expected value of an indicator is a *conditional* probability.

Double expectation

$$E\left(X\right) = E\left(E[X|Y]\right) = E(g(Y))$$

Showing $E(X) = E(E[X|Y])$

Again note $E(E[X|Y])$ is an example of $E(g(Y)$

$$
\begin{aligned}
E(E[X|Y]) &= \int E[X|Y=y] f_y(y)\,dy \\
&= \int \left(\int x f_{x|y}(x|y)\,dx \right) f_y(y)\,dy \\
&= \int \left(\int x \frac{f_{x,y}(x,y)}{f_y(y)}\,dx \right) f_y(y)\,dy \\
&= \int \int x\, f_{x,y}(x,y)\,dx\,dy \\
&= E(h(X,Y)) \\
&= E(X)
\end{aligned}
$$

Double expectation: $E\left(g(X)\right) = E\left(E[g(X)|Y]\right)$

$$E\left(E[I_A(X)|Y]\right) = E[I_A(X)] = Pr\{X \in A\}, \text{ so}$$

$$\begin{aligned} Pr\{X \in A\} &= E\left(E[I_A(X)|Y]\right) \\ &= E\left(Pr\{X \in A|Y\}\right) \\ &= \int_{-\infty}^{\infty} Pr\{X \in A|Y=y\} f_y(y)\, dy, \text{ or} \\ & \quad \sum_y Pr\{X \in A|Y=y\} p_y(y) \end{aligned}$$

This is known as the *Law of Total Probability*

Random Explanatory Variables

Don't you think its strange?

- In the general linear regression model, the $\mathbf{X}$ matrix is supposed to be full of fixed constants.
- This is convenient mathematically. Think of $E(\widehat{\boldsymbol{\beta}})$.
- But in any non-experimental study, if you selected another sample you'd get different $\mathbf{X}$ values, because of random sampling.
- So $\mathbf{X}$ should be at least partly random variables, not fixed.
- View the usual model as *conditional* on $\mathcal{X} = \mathbf{X}$.
- All the probabilities and expected values so far in this course are *conditional* probabilities and *conditional* expected values.
- Conditional on $\mathcal{X} = \mathbf{X}$.
- We don't want to stop there.

$\widehat{\boldsymbol{\beta}}$ is (conditionally) unbiased

$$E(\widehat{\boldsymbol{\beta}}|\mathcal{X} = \mathbf{X}) = \boldsymbol{\beta}$$

For *any* fixed **X** with linearly independent columns.

It's *unconditionally* unbiased too.

$$E\{\widehat{\boldsymbol{\beta}}\} = E\{E\{\widehat{\boldsymbol{\beta}}|\mathcal{X}\}\} = E\{\boldsymbol{\beta}\} = \boldsymbol{\beta}$$

Perhaps Clearer

$$
\begin{aligned}
E\{\widehat{\boldsymbol{\beta}}\} &= E\{E\{\widehat{\boldsymbol{\beta}}|\mathcal{X}\}\} \\
&= \int \cdots \int E\{\widehat{\boldsymbol{\beta}}|\mathcal{X} = \mathbf{X}\}\, f(\mathbf{X})\, d\mathbf{X} \\
&= \int \cdots \int \boldsymbol{\beta}\, f(\mathbf{X})\, d\mathbf{X} \\
&= \boldsymbol{\beta} \int \cdots \int f(\mathbf{X})\, d\mathbf{X} \\
&= \boldsymbol{\beta} \cdot 1 = \boldsymbol{\beta}.
\end{aligned}
$$

Conditional size α test, Critical value f_α

$$Pr\{F > f_\alpha | \mathcal{X} = \mathbf{X}\} = \alpha$$

$$\begin{aligned} Pr\{F > f_\alpha\} &= \int \cdots \int Pr\{F > f_\alpha | \mathcal{X} = \mathbf{X}\} f(\mathbf{X})\, d\mathbf{X} \\ &= \int \cdots \int \alpha f(\mathbf{X})\, d\mathbf{X} \\ &= \alpha \int \cdots \int f(\mathbf{X})\, d\mathbf{X} \\ &= \alpha \end{aligned}$$

A similar calculation applies to confidence intervals and prediction intervals.

The moral of the story

- Don't worry.
- Even though the independent variables are often random, we can apply the usual fixed $\mathbf{X}$ model without fear.
- Estimators are still unbiased.
- Tests have the right Type I error probability.
- Confidence intervals and prediction intervals are still correct.
- And it's all distribution-free with respect to $\mathbf{X}$.

Copyright Information