

Random Independent variables¹

STA302 Fall 2017

¹See last slide for copyright information.

Overview

- 1 Preparation
- 2 Random Independent Variables

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 $I(x \in A)$

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

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

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

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

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

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

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

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

This is known as the *Law of Total Probability*

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 $E(\hat{\beta})$.
- 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 $\mathbf{X} = \mathbf{x}$.
- All the probabilities and expected values so far in this course are *conditional* probabilities and *conditional* expected values.
- Does this make sense?

$\hat{\beta}$ is (conditionally) unbiased

$$E(\hat{\beta} | \mathbf{X} = \mathbf{x}) = \beta \text{ for any fixed } \mathbf{x}$$

It's *unconditionally* unbiased too.

$$E\{\hat{\beta}\} = E\{E\{\hat{\beta} | \mathbf{X}\}\} = E\{\beta\} = \beta$$

Perhaps Clearer

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

Conditional size α test, Critical region A

$$Pr\{F \in A | \mathbf{X} = \mathbf{x}\} = \alpha$$

$$\begin{aligned} Pr\{F \in A\} &= \int \cdots \int Pr\{F \in A | \mathbf{X} = \mathbf{x}\} f(\mathbf{x}) d\mathbf{x} \\ &= \int \cdots \int \alpha f(\mathbf{x}) d\mathbf{x} \\ &= \alpha \int \cdots \int f(\mathbf{x}) d\mathbf{x} \\ &= \alpha \end{aligned}$$

The moral of the story

- Don't worry.
- Even though X variables are often random, we can apply the usual fixed- x model without fear.
- Estimators are still unbiased.
- Tests have the right Type I error probability.
- Similar arguments apply to confidence intervals and prediction intervals.
- And it's all distribution-free with respect to X .

Copyright Information

This slide show was prepared by **Jerry Brunner**, Department of Statistics, University of Toronto. It is licensed under a **Creative Commons Attribution - ShareAlike 3.0 Unported License**. Use any part of it as you like and share the result freely. The \LaTeX source code is available from the course website:
<http://www.utstat.toronto.edu/~brunner/oldclass/302f17>