Probability and Statistics I

22. Conditional Distributions

5.3 Conditional Distributions

Discrete Random Variables

Conditional Probability Mass Function

Let X and Y be two discrete random variables with joint pmf $p(x, y)$. The conditional pmf of Y given $X = x$ $(Y|X = x)$ for any value x such that $p_X(x) > 0$ is

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)}.$$

Conditional Mean and Variance

The expected value and variance of $Y|X = x$ are

$$\mu_{Y|X=x} = E(Y|X = x) = \sum_{y \in D_y} y p_{Y|X}(y|x)$$

and

$$\begin{aligned}\sigma^2_{Y|X=x} = V(Y|X = x) &= E[(Y - E(Y|X = x))^2|X = x] \\ &= E[Y^2|X = x] - E[Y|X = x]^2.\end{aligned}$$

Independence

Two discrete random variables X and Y are said to be independent if

$$p(x, y) = p(x)p(y)$$

for every $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Alternatively, X and Y are independent if

$$p_{Y|X}(y|x) = p_Y(y) \Leftrightarrow p_{X|Y}(x|y) = p_X(x).$$

Example 22.1: Berkeley Admissions Data Revisited

In Example 20.1 we approximated the joint and marginal pmfs of gender (X) and department (Y) of an applicant to as:

Major	Men	Women	$p_Y(y)$
A	0.182	0.024	.206
B	0.124	0.006	.130
C	0.072	0.131	.203
D	0.092	0.083	.175
E	0.042	0.087	.129
F	0.082	0.075	.158
$p_X(x)$	.595	.405	

Example 22.1 ctd

a) Find the conditional pmf of X
   i. given $Y = 1$
   ii. given $Y = 2$.

b) Compute the conditional mean and variance of X
   i. given $Y = 1$
   ii. given $Y = 2$.

c) Comment on the results.

5.3 Conditional Distributions

Continuous Random Variables

Conditional Probability Density Function

Let X and Y be two discrete random variables with joint pdf $f(x, y)$. The conditional pmf of Y given $X = x$ $(Y|X = x)$ for any value x such that $f_X(x) > 0$ is

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}.$$

Conditional Mean and Variance

The expected value and variance of $Y|X = x$ are

$$\mu_{Y|X=x} = E(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\, dy$$

and

$$\begin{aligned}\sigma^2_{Y|X=x} = V(Y|X = x) &= E[(Y - E(Y|X = x))^2|X = x] \\ &= E[Y^2|X = x] - E[Y|X = x]^2.\end{aligned}$$

Independence

Two continuous random variables X and Y are said to be independent if

$$f(x,y) = f_X(x)f_Y(y)$$

for every $(x,y) \in \mathbb{R}^2$.

Alternatively, X and Y are independent if

$$f_{Y|X}(y|x) = f_Y(y) \Leftrightarrow f_{X|Y}(x|y) = f_X(x).$$

Example 22.2

In Example 20.2 we considered the random variables X and Y with joint pdf

$$f(x,y) = \begin{cases} \frac{9}{16}(1-x^2)(1-y^2) & -1 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Compute the conditional pdf of Y given $X = 0$ and $X = .5$.

b) Compute the conditional mean of Y given $X = 0$ and $X = .5$.

c) Comment on the results.

5.3 Conditional Distributions

Bivariate Normal

Bivariate Normal

We say that X and Y have a bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ if

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\frac{\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)}\right).$$

Bivariate Normal

Basic Properties

$$E(X) = \mu_1, V(X) = \sigma_1^2 \qquad E(Y) = \mu_2, V(Y) = \sigma_2^2$$

$$\text{Cov}(X, Y) = \rho\sigma_1\sigma_2 \qquad \text{Corr}(X, Y) = \rho$$

$$E(Y|X = x) = \mu_2 + \rho\sigma_2\frac{x - \mu_1}{\sigma_1} \qquad V(Y|X = x) = \sigma_2^2(1 - \rho^2)$$

X and Y are independent if and only if $\rho = 0$

Bivariate Normal

Marginal Distributions If X and Y have a bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ then the marginal distributions of X and Y are

$$X \sim \text{Normal}(\mu_1, \sigma_1^2)$$

and

$$X \sim \text{Normal}(\mu_2, \sigma_2^2).$$

Bivariate Normal

Conditional Distributions If X and Y have a bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ then the conditional distributions of $X|Y = y$ and $Y|X = x$ are

$$X|Y = y \sim \text{Normal}(E(X|Y = y), V(X|Y = y))$$

and

$$Y|X = x \sim \text{Normal}(E(Y|X = x), V(Y|X = x))$$

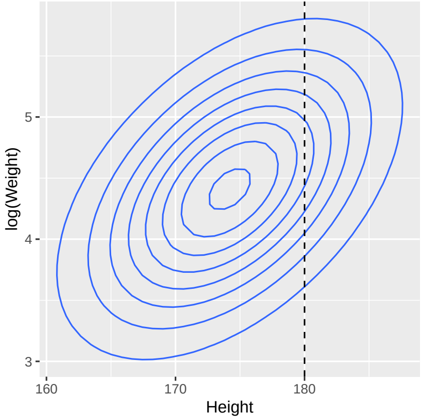
Example 22.3

A study of 646 men between the ages of 50 and 80 conducted by Burmaster and Murray (1998) estimated that the joint distribution of the height (cm), X, and the log of weight (kg), Y, was approximately bivariate normal.
They estimated the means and variances to be

$$\begin{array}{ll} E(X) = 174.20 & E(Y) = 4.41 \\ V(X) = 42.36 & V(Y) = .46 \end{array}$$

with correlation $\rho(X, Y) = .49$.

Example 22.3 ctd



Example 22.3 ctd

a) What is the joint pdf of X and Y?
b) Are height and the log of weight independent?
c) What is the marginal distribution of height?
d) What is the conditional distribution of the log of weight for a man 180 cm tall?
e) Is it unusual for a 180 cm tall man to weigh 70 kg?

5.3 Conditional Distributions

Mean and Variance via
the Conditional Mean and Variance

Conditional Mean and Variance Formulas

For any two random variables, X and Y,

$$E(Y) = E[E(Y|X)]$$
$$V(Y) = E[V(Y|X)] + V[E(Y|X)].$$

Conditional Mean and Variance Formulas

More explicitly, for any two random variables, X and Y,

$$E(Y) = E_X[E_{Y|X}(Y|X)]$$
$$V(Y) = E_X[V_{Y|X}(Y|X)] + V_X[E_{Y|X}(Y|X)].$$

Example 22.3 ctd

a) What is the joint pdf of X and Y?
b) Are height and the log of weight independent?
c) What is the marginal distributions of height?
d) What is the conditional distribution of the log of weight for a man 180 cm tall?
e) Is it unusual for a 180 cm tall man to weigh 70 kg?
f) Verify the formulas for computing the mean and variance of Y from the conditional mean and variance formulas for the case of two bivariate normal random variables.

Questions?

T

he daytime mean temperature for London in November is approximately normally distributed with a mean of 3.4 C and standard deviation of 1.7 C. Suppose that the joint distribution of the temperature subsequent days is bivariate normal with a correlation of .6.

Let T_1 and T_2 be the temperature on two different days.

1. Sketch a plot showing contours of the joint pdf of T_1 and T_2.
2. What is the distribution of T_2 given $T_1 = 1$ C? Be as specific as possible.
3. Explain the result in terms of the phenomenon of regression to the mean.