743- Regression and Time Series

Mamikon S. Ginovyan

Conditional Probabilities & Expectations

1. Conditional Probabilities

- Assume that we have two **DRV's** ***X*** and ***Y*** with **joint pmf** $f(x,y) = P(X = x, Y = y)$ and **marginal pmf's** $f_X(x) = P(X = x)$ and $f_Y(y) = P(Y = y)$, respectively.

- **Def.** The **conditional pmf's** $f(y|x)$ and $f(x|y)$ are defined to be

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

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

Conditional Probabilities

Observe that

$$f(x,y) = f(x\,|\,y) f_Y(y) = f(y\,|\,x) f_X(x), \quad \text{and}$$

$$\sum_y f(y\,|\,x) = \sum_x f(x\,|\,y) = 1.$$

Similarly, if ***X*** and ***Y*** are **CRV's** with **joint *pdf*** $f(x,y)$, and **marginal *pdf*'s** $f_X(x)$ and $f_Y(y)$, then the **conditional *pdf*'s** $f(x|y)$ and $f(y|x)$ are defined to be

$$f(y\,|\,x) = \frac{f(x,y)}{f_X(x)} \quad \text{and} \quad f(x\,|\,y) = \frac{f(x,y)}{f_Y(y)}.$$

Conditional Probabilities

Again we have $f(x,y) = f(x\,|\,y)f_Y(y) = f(y\,|\,x)f_X(x),$ and

$$\int_{-\infty}^{\infty} f(y\,|\,x)dy = \int_{-\infty}^{\infty} f(x\,|\,y)dx = 1.$$

For **conditional *cdf*** we have

$$F(x\,|\,y) = P(X \le x\,|\,Y = y) = \int_{-\infty}^{x} f(t\,|\,y)dt \qquad \text{for } \mathbf{CRV},$$

$$= \sum_{t \le x} f(t\,|\,y) \qquad \text{for } \mathbf{DRV}.$$

2. Conditional Expectations

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two RV's defined on the same probability space $(\Omega, \Im, P)$, with **joint, marginal,** and **conditional** $f(x, y)$, $f_X(x)$, $f_Y(y)$, $f(y\,|\,x)$ *and* $f(x\,|\,y)$ **pdf**'s (or **pmf**'s), respectively.

•**Def.** The **conditional expectation** of $\boldsymbol{Y}$ given that $\boldsymbol{X} = x$ is a **function** of x, and is given by

$$y(x) = E[Y\,|\,X = x] = \int_{-\infty}^{\infty} y f(y\,|\,x)dy \quad \text{for } \textbf{CRV},$$

$$= \sum_{y} y f(y\,|\,x) \quad \text{for } \textbf{DRV}.$$

Conditional Expectations

- **Remark 1.** The conditional expectation has all of the proportions of the ordinary expectation:

- Let ***X*** and ***Y*** be two RV's and let c_1, c_2 and c_3 be constants. Then for any functions $g_1(x)$ and $g_2(x)$ for which $E[g_1(x)]$ and $E[g_2(x)]$ exist, we have:

(a) $$E[c_1 g_1(X) + c_2 g_2(X) + c_3 \mid Y = y]$$
$$= c_1 E[g_1(X) \mid Y = y] + c_2 E[g_2(X) \mid Y = y] + c_3.$$

(b) If $g_1(x) \geq 0$ for all $\boldsymbol{x}$, then

$$E[g_1(X) \mid Y = y] \geq 0.$$

Conditional Expectations

(c) If $g_1(x) \geq g_2(x)$ for all $\boldsymbol{x}$, then

$$E[g_1(X) \mid Y = y] \geq E[g_2(X) \mid Y = y].$$

(d) If $c_1 \leq g_1(x) \leq c_2$ for all $\boldsymbol{x}$, then

$$c_1 \leq E[g_1(X) \mid Y = y] \leq c_2.$$

Conditional Expectations

Remark 2 (Prediction Problem). Let X and Y be two random variables with finite second moments.

We will prove that:

- The random variable $Z = y(X) = E(Y|X)$, which is a function of X, provides the **best "guess"** of Y based on knowledge of X:

More precisely,

(a) $$\min_{g(x)} E(Y - g(X))^2 = E(Y - E(Y|X))^2.$$

(b) In the special case of (a) with $X = x =$ **constant**, we have

$$\min_{b} E(X - b)^2 = E(X - EX)^2.$$

Conditional Expectations

- **Conditional Expectation of a function of a RV.**

Def.1. If $Z=g(X)$ is a function of $\boldsymbol{X}$, then the **conditional expectation** of $\boldsymbol{Z}$ given that $\boldsymbol{Y=y}$ is given by

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

$$=\int_{-\infty}^{\infty} g(x)f(x\,|\,y)dx, \qquad \text{for } \mathbf{CRV}, \qquad (1\boldsymbol{a})$$

$$=\sum_{y} g(x)f(x\,|\,y), \qquad \text{for } \mathbf{DRV}. \qquad (1\boldsymbol{b})$$

Computing by Conditioning

I. Computing Expectations by Conditioning
II. Computing Variances by Conditioning
III. Computing Probabilities by Conditioning

Computing by Conditioning

I. Computing Expectations by Conditioning

Conditional Expectation as a RV.

So far we have considered the conditional expectation $E[Y \mid X = x]$ for a **fixed** value $\boldsymbol{X = x}$.

- Denote by $E[Y \mid X]$ that function of the RV $\boldsymbol{X}$**,** whose value at $\boldsymbol{X = x}$ is $E[Y \mid X = x]$.

- Note that $E[Y \mid X]$, as a function of RV $\boldsymbol{X}$, is **itself a RV** with pdf (or pmf) $f_X(x)$.

Computing Expectations by Conditioning

An extremely important property of conditional expectations as a RV is given by the following.

Theorem 1. For any two RV's $\boldsymbol{X}$ and $\boldsymbol{Y}$

$$E(Y) = E\left[E(Y \mid X)\right], \qquad (1)$$

provided that the expectations exist.

Computing Expectations by Conditioning

Proof for DRV's X and Y.

We will use the formula

$$f(x,y) = f(x\mid y)f_Y(y) = f(y\mid x)f_X(x).$$

We have

$$E[E(Y\mid X)] = \sum_{x\in D_X} E(Y\mid X=x)f_X(x)$$

$$= \sum_{x\in D_X}\left[\sum_{y\in D_Y} y f_{Y|X}(y\mid x)\right]f_X(x)$$

$$= \sum_{x\in D_X}\sum_{y\in D_Y} y\frac{f(x,y)}{f_X(x)}f_X(x) = \sum_{y\in D_Y} y\left[\sum_{x\in D_X} f(x,y)\right]$$

$$= \sum_{y\in D_Y} y f_Y(y) = E(Y)$$

Computing Variances by Conditioning

II. Computing Variances by Conditioning

Conditional variance formula

Def. The **conditional variance** of $\boldsymbol{X}$ given that $\boldsymbol{Y=y}$ is defined by

$$Var(X \mid Y=y) = E\left[\left(X - E\left[X \mid Y=y\right]\right)^2 \mid Y=y\right]$$

$$E\left[X^2 \mid Y=y\right] - \left(E\left[X \mid Y=y\right]\right)^2.$$

That is, the conditional variance is defined in exactly the same manner as the ordinary variance with the exception that all probabilities are determined conditional on the event that $\boldsymbol{Y=y}$.

Computing Variances by Conditioning

Conditional Variance as a RV

Letting $Var(X \mid Y)$ denote the RV (which is a function of $\boldsymbol{Y}$), whose value when $\boldsymbol{Y=y}$ is $Var(X \mid Y = y)$, we have the following important result.

Theorem 2. (The Conditional Variance Formula).
For any two RV's $\boldsymbol{X}$ and $\boldsymbol{Y}$

$$Var(Y) = Var\left[E(Y \mid X)\right] + E\left[Var(Y \mid X)\right], \qquad (1)$$

provided that the expectations exist.

Computing Variances by Conditioning

Corollary (Conditional Variance Inequality)

Since the second term in (1) is non-negative, it is expectation of a non-negative RV:

$$E\left[Var(Y \mid X)\right] = E\left(Y - E(Y \mid X)\right)^2 \geq 0$$

from (1) we obtain the

Conditional Variance Inequality: For any two RV's ***X*** and ***Y***

$$Var(E\left[Y \mid X\right]) \leq Var\left(Y\right).$$

Computing Probabilities by Conditioning

II. Computing Probabilities by Conditioning

Now we show that, **conditioning on an appropriate RV**, can be used to compute probabilities. To see this, for an event ***A*** consider the indicator RV:

$$X = I_A = \begin{cases} 1, & \text{if } \; A \; \text{ occurs} \\ 0, & \text{if } \; A^c \; \text{ occurs.} \end{cases}$$

We know that $E[X] = P(A)$. Hence, for any RV ***Y***,

$$E[X \mid Y = y] = P(A \mid Y = y).$$

Computing Probabilities by Conditioning

Therefore, by Theorem 1: $(E[X] = E[E[X \mid Y]])$,

we obtain

$$P(A) = E[X] = E[E[X \mid Y]]$$

$$= \sum_y E[X \mid Y = y] f_Y(y) = \sum_y P(A \mid Y = y) f_Y(y), \quad \text{if } \boldsymbol{Y} \text{ is } \textbf{DRV},$$

$$= \int_{-\infty}^{\infty} E[X \mid Y = y] f_Y(y) dy = \int_{-\infty}^{\infty} P(A \mid Y = y) f_Y(y) dy, \quad \text{if } \boldsymbol{Y} \text{ is } \textbf{CRV}.$$

Examples

- **Example 1.** The joint *pdf* of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is given by

$$f(x,y) = \begin{cases} \frac{1}{2} y e^{-xy}, & 0 < x < \infty, \quad 0 < y < 2, \\ 0, & \text{otherwise.} \end{cases}$$

What is $E\left[e^{X/2} \mid Y = 1\right]$?

Example 1.-Solution

Solution: The conditional *pdf* of $\boldsymbol{X}$, given that $\boldsymbol{Y = 1}$, is given by

$$f_{X|Y}(x\,|\,1) = \frac{f(x,1)}{f_Y(1)} = \frac{\frac{1}{2}e^{-x}}{\int_0^\infty \frac{1}{2}e^{-x}dx} = e^{-x},$$

because $\int_0^\infty e^{-x}dx = -e^{-x}\Big|_0^\infty = 1.$

Hence, $E\left[e^{X/2}\,|\,Y = 1\right] = \int_0^\infty e^{x/2} f_{X|Y}(x\,|\,1)dx$

$$= \int_0^\infty e^{x/2}e^{-x}dx = \int_0^\infty e^{-x/2}dx = 2.$$

Examples

Example 2. An insurance company supposes that the number of accidents that each of its policyholders will have in a year is **Poisson distributed**, with the mean of the Poisson depending on the policyholder. If the Poisson mean Λ of a randomly chosen policyholder has a gamma $\Lambda \sim G(a = 2, b = 1)$ distribution with pdf

$$g_\Lambda(l) = l e^{-l}, l \geq 0,$$

what is the probability that a randomly chosen policyholder has **exactly** $\boldsymbol{n}$ accidents next year?

Example 2-Preliminaries

Preliminaries:

- The **Gamma function**

$$\Gamma(a) = \int_0^\infty y^{a-1} e^{-y} dy, a > 0.$$

1. $\Gamma(1) = 1.$

2. $\Gamma(a+1) = a \cdot \Gamma(a), a > 0.$

3. If a is an integer, then $\Gamma(a+1) = a!.$
   $\Gamma(a)$ is an extension of the factorial.

Example 2-Preliminaries

- The **Gamma distribution**:

A RV $X \sim G(a, b), a > 0, b > 0,$

If the *pdf* *f(x)* is given by

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(a)b^{a}} \cdot x^{a-1} e^{-x/b}, & 0 < x < \infty \\ 0, & \text{elsewhere.} \end{cases}$$

Example 2-Solution

Solution: Let $\boldsymbol{X}$ denote the number of accidents that a randomly chosen policyholder has next year.

Letting $\boldsymbol{Y} = \Lambda$ be the Poisson mean number of accidents for this policyholder, conditioning on $\boldsymbol{Y}$ yields

$$P\{X = n\} = \int_0^\infty P\{X = n \mid Y = l\} g(l) d l$$

$$= \int_0^\infty \left(\frac{e^{-l} l^n}{n!} \right) l e^{-l} d l = \frac{1}{n!} \int_0^\infty l^{n+1} e^{-2l} d l \qquad (1)$$

Example 2-Solution

To compute the integral, observe that the function

$$h(l) = \frac{e^{-2l}\, l^{\,n+1}}{\Gamma(n+2)(1/2)^{(n+2)}} = \frac{2^{(n+2)} e^{-2l}\, l^{\,n+1}}{(n+1)!}, \quad l > 0$$

is the *pdf* of gamma $G(n+2, 1/2) - RV$, and hence

$$1 = \int_0^\infty h(l)\, dl = \int_0^\infty \frac{2^{(n+2)} e^{-2l}\, l^{\,n+1}}{(n+1)!}\, dl$$

$$= \frac{2^{n+2}}{(n+1)!} \int_0^\infty l^{\,n+1} e^{-2l}\, dl\,.$$

Example 2-Solution

Therefore,

$$\int_0^\infty l^{\,n+1} e^{-2l}\, dl = \frac{(n+1)!}{2^{n+2}}. \qquad (2)$$

From (1) and (2), we conclude

$$P\{X = n\} = \frac{1}{n!}\frac{(n+1)!}{2^{n+2}} = \frac{n+1}{2^{n+2}}.$$

Remark:

$$P\{X = 1\} = \frac{1}{4} = .25; \quad P\{X = 2\} = \frac{3}{16} \approx .19.$$

Examples

- **Example 3.** Suppose that **X** and $\boldsymbol{Y}$ are **independent** continuous RV's having pdf's $f_X(x)$ and $f_Y(y)$, respectively.

Compute:

$(a)\quad P\{X < Y\} = \{w : X(w) < Y(w)\}.$

(b) The *pdf* of the RV $\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{Y}$.

Example 3.-Solution

- **Solution**: **(a)** Conditioning on the value of $\boldsymbol{Y}$ yields

$$P\{X < Y\} = \int_{-\infty}^{\infty} P\{X < Y \mid Y = y\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} P\{X < y \mid Y = y\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} P\{X < y\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} F_X(y) f_Y(y)dy,$$

where $\quad F_X(y) = \int_{-\infty}^{y} f_X(x)dx.$

Example 3.-Solution

- **Question**.

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be **independent**, and $\boldsymbol{X} \sim U(0,1)$, $\mathrm{Y} \sim U(0,1)$.

What is $\mathrm{P}\{\boldsymbol{X} < \boldsymbol{Y}\}$?

Example 3.-Solution

- **Question.**

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be **independent**, and $\boldsymbol{X} \sim U\,(0,1)$, Y $\sim U\,(0,1)$.
What is P{ $\boldsymbol{X} < \boldsymbol{Y}$ } ?

We have $f_Y(y) = I_{(0,1)}(y)$ and $F_X(y) = y, 0 < y < 1.$

Hence $$P\{X < Y\} = \int_{-\infty}^{\infty} F_X(y) f_Y(y)\,dy$$

$$= \int_0^1 (y \cdot 1)\,dy = \left.\frac{y^2}{2}\right|_0^1 = \frac{1}{2}.$$

The answer in this case is intuitively obvious because of the symmetry of the problem in $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Example 3.-Solution

(b) First we compute the cdf $F_Z(z)$ of $Z = X + Y$.

Conditioning on the value of $\boldsymbol{Y}$ yields

$$F_Z(z) = P\{X+Y \le z\} = \int_{-\infty}^{\infty} P\{X+Y \le z \mid Y = y\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} P\{X \le z-y \mid Y = y\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} P\{X \le z-y\} f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} F_X(z-y) f_Y(y)dy, \qquad (1)$$

where $F_X(z-y) = \int_{-\infty}^{z-y} f_X(x)dx.$

Example 3.-Solution

If we differentiate (1) w.r.t. $\mathbf{z}$, then we get the *pdf* $f_Z(z)$ of $\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{Y}$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy \qquad (2)$$

•**Remark.** Formula (2) is called the **convolution** of the functions $f_X(x)$ and $f_Y(y)$.

Mixture Models

Def. We say that a RV X has **mixture distribution** if the distribution of X depends on a quantity that also has a distribution, that is, the distribution of X depends on some RV Y.

- In this case we say that we have a **Mixture Model** $X \sim Y$.

Examples

- **Example 4.** Let X_1 and X_2 be RV's such that

$$X_1 \mid X_2 \sim Bin(X_2, p), \qquad (\text{so } E[X_1 \mid X_2] = X_2 \cdot p \ ); \qquad (1)$$

and

$$X_2 \sim Poisson(l), \qquad (\text{so } E[X_2] = l \ ); \qquad (2)$$

Find $E[X_1]$.

Example 4.- Solution

- **Solution.** By Th. 1

$$(E(Y) = E[E(Y \mid X)]),$$

we have

$$E[X_1] = E\left(E[X_1 \mid X_2]\right) \qquad by \quad (1)$$

$$= E[p \cdot X_2] = p \cdot E[X_2] \qquad by \quad (2)$$

$$= p \cdot l.$$

Thus, $E[X_1] = p \cdot l.$

Examples

- **Example 5.** Let X_1, X_2, X_3 be RV's such that

$$X_1 \mid X_2 \sim Bin(X_2, p), \qquad (\text{so } E[X_1 \mid X_2] = X_2 \cdot p \ ); \qquad (1)$$

$$X_2 \mid X_3 \sim Poisson(X_3), \qquad (\text{so } E[X_2 \mid X_3] = X_3 \ ); \qquad (2)$$

$$X_3 \sim Exp(b), \qquad (\text{so } E[X_3] = b \ ); \qquad (3)$$

Find $E[X_1]$.

Example 5.-Solution

- **Solution**. By Th. 1, we have

$$E[X_1] = E(E[X_1 | X_2]) \qquad by \quad (1)$$

$$= E[p \cdot X_2] \qquad by \quad Th.1$$

$$= E(E[p \cdot X_2 | X_3]) \qquad by \quad (2)$$

$$= E[p \cdot X_3] = p \cdot E[X_3] \qquad by \quad (3)$$

$$= p \cdot b.$$

- **Thus,** $E[X_1] = p \cdot b.$

Summary Example.

A large insurance agency services a number of customers who have purchased both ***an automobile policy (AP)*** and ***a homeowner's policy (HP)*** from the agency.

For each type of policy, a **deductible amount** must be specified.

For AP, the choices are \$100 and \$250,

For HP, the choices are \$0, \$100, and \$200.

- Suppose an individual with **both types** of policy is selected at random from the agency's files.

Summary Example.

- Let X = the deductible amount on the AP, and
  Y = the deductible amount on the HP.
- Possible pairs are then

(100, 0), (100,100), (100, 200), (250, 0) (250, 100), and (250, 200).

The **joint pmf** specifies the probabilities associated with each one of these pairs, with any other pair having probability zero.

- Suppose the **joint pmf** is given in the accompanying **joint probability table:**

Summary Example.

$p(x,y)=P(X=x, Y=y)$		y		
		0	100	200
x	100	*.20*	*.10*	*.20*
	250	*.05*	*.15*	*.30*

Summary Example

(1) **Compute the marginal pmf's** $p_X(x) = P(X = x)$ and $p_Y(y) = P(Y = y)$.

- **Solution.**

The possible values of X are x **=100** and x **=250** , so computing row totals in the joint probability table yields

$$Px(100) = p(100,0) + p(100,100) + p(100,200) = .50,$$

and

$$Px(250) = p(250,0) + p(250,100) + p(250,200) = .50.$$

Summary Example

- Thus, the **marginal *pmf*** $p_X(x)$ of $\boldsymbol{X}$ is then

$$p_X(x) = \begin{cases} .5 & x = 100, 250 \\ 0 & otherwise. \end{cases}$$

- Similarly, the **marginal *pmf*** $p_Y(y)$ of $\boldsymbol{Y}$ is obtained

from column totals as

$$p_Y(y) = \begin{cases} .25 & y = 0, 100 \\ .50 & y = 200 \\ 0 & \text{otherwise} \end{cases} .$$

Summary Example

(2) Find marginal distributions (tables) of X and Y.

•**Solution.**

The **marginal distributions** of X and Y are given by

x	*100*	*250*
$P_X(x)$	$.5$	$.5$

y	*0*	*100*	*200*
$P_Y(y)$	$.25$	$.25$	$.5$

Summary Example

(3) Compute the probability $P(Y \geq 100)$.

•**Solution.**

We have

$$P(Y \geq 100) = p_Y(100) + p_Y(200) = .75.$$

(4) Compute $m_X,\ m_Y,\ s_X^2,\ s_Y^2,\ Cov(X,Y)$ and $r(X,Y)$.

•**Solution.**

We have

$$m_X = E[X] = \sum x p_X(x) = 175,$$

$$m_Y = E[Y] = \sum y p_Y(y) = 125.$$

Summary Example

$$s_X^2 = E[X^2] - (E[X])^2 = 36250 - (175)^2 = 5625, \quad s_X = 75,$$

$$s_Y^2 = E[Y^2] - (E[Y])^2 = 22500 - (125)^2 = 6875, \quad s_Y = 82.92.$$

$$Cov(X,Y) = \sum_x \sum_y (x-175)(y-125)\,p(x,y)$$

$$= (100-175)(0-125)(.20) + \cdots + (250-175)(200-125)(.30)$$

$$= 1875.$$

$$r(X,Y) = \frac{Cov(X,Y)}{s_X s_Y} = \frac{1875}{(75)(82.92)} = .301.$$

Summary Example

(5) Compute the conditional *p.m.f*'s $p_{Y|X}(y\,|\,x)$ and $p_{X|Y}(x\,|\,y)$.

•**Solution**.

For $p_{Y|X}(y\,|\,x)$ we have

$$p_{Y|X}(200\,|\,250) = \frac{p(250,200)}{p_X(250)} = \frac{.3}{.5} = .6,$$

$$p_{Y|X}(0\,|\,250) = \frac{p(250,0)}{p_X(250)} = \frac{.05}{.5} = .1,$$

$$p_{Y|X}(100\,|\,250) = \frac{p(250,100)}{p_X(250)} = \frac{.15}{.5} = .3.$$

Summary Example

Thus,

$$p_{Y|X}(0\,|\,250)+p_{Y|X}(100\,|\,250)+p_{Y|X}(200\,|\,250)=.1+.3+.6=1.$$

For $p_{X|Y}(x\,|\,y)$ we have

$$p_{X|Y}(100\,|\,0)=\frac{p(100,0)}{p_Y(0)}=\frac{.20}{.25}=.8,$$

$$p_{X|Y}(250\,|\,0)=\frac{p(250,0)}{p_Y(0)}=\frac{.05}{.25}=.2.$$

Again, the conditional probabilities add to 1.

Summary Example

(6) Compute the conditional expectations and variances

$$m_{Y|X=x} = E(Y \mid X = x), \quad s^2_{Y|X=x} = V(Y \mid X = x) \quad \text{and}$$

$$m_{X|Y=y} = E(X \mid Y = y), \quad s^2_{X|Y=y} = V(X \mid Y = y).$$

Solution. Using the conditional distributions found in Part (5), we have

For $m_{Y|X=x} = E(Y \mid X = x),$

$$m_{Y|X=250} = E(Y \mid X = 250)$$

$$= (0)\, p_{Y|X}(0 \mid 250) + (100)\, p_{Y|X}(100 \mid 250) + (200)\, p_{Y|X}(200 \mid 250)$$

$$= 0(.1) + 100(.3) + 200(.6) = 150.$$

Summary Example

Remark. Given that the possibilities for Y are 0, 100, and 200 and most of the probability is on 100 and 200, it is reasonable that the conditional mean should be between 100 and 200.

For $s^2_{Y|X=x} = V(Y \mid X = x)$,

$$E(Y^2 \mid X = 250)$$

$$= 0^2 p_{Y|X}(0 \mid 250) + 100^2 p_{Y|X}(100 \mid 250) + 200^2 P_{Y|X}(200 \mid 250)$$

$$= 0^2(.1) + 100^2(.3) + 200^2(.6) = 27{,}000.$$

Thus,

$$s^2_{Y|X=250} = V(Y \mid X = 250) = E(Y^2 \mid X = 250) - m^2_{Y|X=250}$$

$$= 27{,}000 - 150^2 = 4500.$$

Summary Example-Remark

- Taking the square root, we get $s_{Y|X=x} = 67.08$,

- **Remark.** It is important to realize that $E(Y^2 \mid X = x)$ is one particular possible value of the RV $E(Y^2 \mid X)$ which is a function of $\boldsymbol{X}$.

- **Remark.** Similarly, the conditional variance $Var(Y \mid X = x)$ is a value of the RV $Var(Y \mid X)$.

Summary Example-Remark

The value of $\boldsymbol{X}$ might be 100 or 250.

So far we have just $E(Y \mid X = 250) = 150$ and

$Var(Y \mid X = 250) = 4500.$

Similarly for $\boldsymbol{X} = 100$ we obtain $E(Y \mid X = 100) = 100$ and

$Var(Y \mid X = 100) = 8000.$

Here is a summary table of the obtained results.

x	$P(X=x)$	$E(Y \mid X=x)$	$Var(Y \mid X=x)$
100	**.5**	**100**	**8000**
250	**.5**	**150**	**4500**

Summary Example-Remark

Similarly, we can compute the conditional mean $\boldsymbol{E}(\boldsymbol{X} \mid \boldsymbol{Y} = y)$ and variance $Var\,(\boldsymbol{X} \mid \boldsymbol{Y} = y)$.

We have

$$m_{X|Y=0} = E(X \mid Y=0) = (100)p_{X|Y}(100 \mid 0) + (250)p_{X|Y}(250 \mid 0)$$
$$= (100)(.8) + (250)(.2) = 130.$$

$$s^2_{X|Y=0} = V(X \mid Y=0) = E([X - E(X \mid Y=0]^2 \mid Y=0)$$
$$= (100-130)^2 p_{X|Y}(100 \mid 0) + (250-130)^2 p_{X|Y}(250 \mid 0)$$
$$= (30)^2(.8) + (120)^2(.2) = 3600.$$

Summary Example-Remark

- Similar calculations give the other entries of the following table:

y	$P(Y=y)$	$E(X \mid Y=y)$	$Var(X \mid Y=y)$
0	.25	130	3600
100	.25	190	5400
200	.5	190	5400

Summary Example

(7) Verify Theorems 1 and 2.

Solution.

For the mean (Theorem 1). $E\left[E(Y \mid X)\right] = E(Y).$

We use the following table for conditional mean and variance, obtained in Part (5):

x	$P(X=x)$	$E(Y \mid X=x)$	$Var(Y \mid X=x)$
100	.5	100	8000
250	.5	150	4500

Summary Example

- **First,** compute

$$E[E(Y|X)]$$
$$= E(Y|X=100)P(X=100) + E(Y|X=250)P(X=250)$$
$$= 100(.5) + 150(.5) = 125.$$

- **Second,** compute $\boldsymbol{E}(\boldsymbol{Y})$ using the marginal distribution of $\boldsymbol{Y}$:

y	*0*	*100*	*200*
$P_Y(y)$	.25	.25	.5

$$E(Y) = (0)P(Y=0) + (100)P(Y=100) + (200)P(Y=200)$$
$$= 0(.25) + 100(.25) + 200(.5) = 125.$$

Summary Example

- **Thus,**

$$E\left[E(Y \mid X)\right] = E(Y) = 125,$$

in agreement with the calculation based on the Theorem 1.

- **For the variance** **(Theorem 2)** .

$$V(Y) = V\left[E(Y \mid X)\right] + E\left[V(Y \mid X)\right].$$

- **First,** compute the mean of the conditional variance:

$$E\left[Var(Y \mid X)\right]$$

$$= Var(Y \mid X = 100)P(X = 100) + V(Y \mid X = 250)P(X = 250)$$

$$= 4500(.5) + 8000(.5) = 6250.$$

Summary Example

- **Second,** compute the variance of the conditional mean.

$$Var\left[E(Y\mid X)\right] = .5(100-125)^2 + .5(150-125)^2 = 625.$$

So

$$Var\left[E(Y\mid X)\right] + E\left[Var(Y\mid X)\right] = 625 + 6250 = 6875.$$

In part (3) we have computed

$$s_Y^2 = Var(Y) = E[Y^2] - (E[Y])^2 = 22500 - (125)^2 = 6875.$$

Thus,

$$Var(Y) = Var\left[E(Y\mid X)\right] + E\left[Var(Y\mid X)\right] = 6875,$$

in agreement with the calculation based on the theorem 2.