Useful facts

$$f_{X,Y}(x,y) = f_X(x)\, f_{Y|X}(y|x) \qquad (1)$$

$$E[E(Y|X)] = EY \quad \text{(switch order)} \qquad (2)$$

$$\operatorname{Var} Y = E[\operatorname{Var}(Y|X)] + \operatorname{Var}[E(Y|X)] \qquad (3)$$

Interpretation of (2) and (3):

$E(Y|X=x)$ is a function of x. (*)

$E(Y|X)$ is a random variable which is a function of the r.v. X

(It is the function (*) evaluated at the random value X.)

$\operatorname{Var}(Y|X=x)$ is a function of x. (* *)

$\operatorname{Var}(Y|X)$ is a random variable which is a function of the r.v. X.

(It is the function (* *) evaluated at the random value X.)

Proof of $EY = E[E(Y|X)]$

(continuous case)

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

$$= \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \right) dy$$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y \underbrace{f_{X,Y}(x,y)}_{f_X(x) f_{Y|X}(y|x)}\,dy \right) dx$$

$$= \int_{-\infty}^{\infty} f_X(x) \underbrace{\left(\int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy \right)}_{E(Y|X=x)} dx$$

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

Proof of (2):

$$\mathrm{Var}(Y|X) = E(Y^2|X) - [E(Y|X)]^2$$

This is just the usual formula for the variance $\mathrm{Var}\,Y = EY^2 - (EY)^2$ applied to the conditional distn. of $Y|X$.

Take expectations:

$$E\,\mathrm{Var}(Y|X) = \underbrace{E\,E(Y^2|X)}_{= EY^2 \text{ by (1) applied to } Y^2.} - E[E(Y|X)]^2$$

Now note that

$$\mathrm{Var}[E(Y|X)] = E[E(Y|X)]^2 - [\underbrace{E\,E(Y|X)}_{= EY \text{ by (1).}}]^2$$

by applying the usual formula for variance to the r.v. $E(Y|X)$.

Adding these together gives

$$E[\mathrm{Var}(Y|X)] + \mathrm{Var}[E(Y|X)]$$
$$= EY^2 - (EY)^2 = \mathrm{Var}\,Y.$$

QED

Example

		X			
	1	1/20	2/20	3/20	4/20
Y	2	0	1/20	2/20	3/20
	3	0	0	1/20	2/20
	4	0	0	0	1/20
		1	2	3	4

$f_X(x)$	X	$E(Y\|X)$	$Var(Y\|X)$
$\frac{1}{20}$	1	1	0
$\frac{3}{20}$	2	4/3	$\frac{2}{9}$
$\frac{6}{20}$	3	5/3	$\frac{5}{9}$
$\frac{10}{20}$	4	2	1

$$EY = E\,E(Y|X) = 1\cdot\frac{1}{20} + \frac{4}{3}\cdot\frac{3}{20} + \frac{5}{3}\cdot\frac{6}{20} + 2\cdot\frac{10}{20}$$

$$= \frac{7}{4}$$

$$var[E(Y|X)] = \left(1-\frac{7}{4}\right)^2\cdot\frac{1}{20} + \left(\frac{4}{3}-\frac{7}{4}\right)^2\cdot\frac{3}{20}$$

$$+ \left(\frac{5}{3}-\frac{7}{4}\right)^2\cdot\frac{6}{20} + \left(2-\frac{7}{4}\right)^2\cdot\frac{10}{20}.$$

$$= \frac{7}{80}$$

Alternatively,

$$\text{Var}\, E(Y|X) = E\left[(E(Y|X))^2\right] - \left(E\left[E(Y|X)\right]\right)^2$$

$$= 1^2 \cdot \frac{1}{20} + \left(\frac{4}{3}\right)^2 \cdot \frac{3}{20} + \left(\frac{5}{3}\right)^2 \cdot \frac{6}{20} + 2^2 \cdot \frac{10}{20} - \left(\frac{7}{4}\right)^2$$

$$= \frac{7}{80}.$$

$$E\, \text{Var}(Y|X) = 0 \cdot \frac{1}{20} + \frac{2}{9} \cdot \frac{3}{20} + \frac{5}{9} \cdot \frac{6}{20} + 1 \cdot \frac{10}{20}$$

$$= \frac{7}{10}$$

Thus $\text{Var}\, Y = \text{Var}\, E(Y|X) + E\, \text{Var}(Y|X)$

$$= \frac{7}{80} + \frac{7}{10} = \frac{63}{80}$$

By direct calculation

(which is much easier in this case)

Marginal for Y :

y	$f_Y(y)$
1	$10/20 = \frac{1}{2}$
2	$6/20 = 3/10$
3	$3/20$
4	$1/20$

$$EY = 1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{10} + 3 \cdot \frac{3}{20} + 4 \cdot \frac{1}{20}$$

$$= 7/4$$

$$\operatorname{Var} Y = 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{3}{10} + 3^2 \cdot \frac{3}{20} + 4^2 \cdot \frac{1}{20} - \left(\frac{7}{4}\right)^2$$

$$= \frac{63}{80}$$

Example

Joint pdf $f_{X,Y}(x,y) = 6(1-x-y)\,I_D(x,y)$

where $D = \{(x,y) : x>0, y>0, x+y<1\}$.

We obtained

$$f_X(x) = 3(1-x)^2 \text{ for } 0<x<1$$

$$E(Y|X=x) = \frac{1-x}{3}, \quad Var(Y|X=x) = \frac{(1-x)^2}{18}.$$

Thus $E(Y|X) = \frac{1-X}{3}$, $Var(Y|X) = \frac{(1-X)^2}{18}$.

$$EY = E\,E(Y|X) = E\left(\frac{1-X}{3}\right)$$

$$= \frac{1}{3} - \frac{1}{3}EX = \frac{1}{4}$$

where $EX = \int_0^1 x\, 3(1-x)^2 dx = \frac{1}{4}$

$$Var(E(Y|X)) = Var\left(\frac{1-X}{3}\right) = \frac{1}{9}Var(1-X)$$

$$= \frac{1}{9}Var(-X)$$

$$= \frac{1}{9}Var(X)$$

__Example__: $f(x,y) = 6(1-x-y)\, I_D(x,y)$

$$\operatorname{Var} Y = E \operatorname{Var}(Y|X) + \operatorname{Var}(E(Y|X))$$

$$\left(\begin{aligned} \operatorname{Var}(Y|X=x) &= \frac{(1-x)^2}{18} \\ E(Y|X=x) &= \frac{1-x}{3} \end{aligned}\right)$$

$$= E\frac{(1-X)^2}{18} + \operatorname{Var}\left(\frac{1-X}{3}\right)$$

$$= \frac{1}{18} E(1-X)^2 + \frac{1}{9} \underbrace{\operatorname{Var}(1-X)}_{\underbrace{\operatorname{Var}(-X)}_{\operatorname{Var} X}}$$

$$\boxed{f_X(x) = 3(1-x)^2, \quad 0<x<1}$$

$$= \frac{1}{18}\left(1 - 2EX + EX^2\right) + \frac{1}{9}\left[EX^2 - (EX)^2\right]$$

$$=$$

$$EX = \frac{1}{4}$$

$$EX^2 = \int_0^1 x^2\, 3(1-x)^2\, dx =$$

$$3B(3,3) = \frac{\cancel{3}\cdot\cancel{2}\cdot 2}{5\cdot 4\cdot \cancel{3}\cdot\cancel{2}} = \frac{1}{10}$$

$$= \frac{1}{18}\left(1 - \frac{2}{4} + \frac{1}{10}\right) + \frac{1}{9}\left[\frac{1}{10} - \frac{1}{16}\right]$$

$$= \frac{1}{18} \cdot \frac{6}{10} + \frac{1}{9} \cdot \underbrace{\frac{16-10}{160}}_{\frac{3}{80}}$$

$$= \frac{1}{30} + \frac{1}{240} = \frac{9}{240} = \frac{3}{80} = \text{Var } Y$$

$$= \frac{1}{18}\left(1-\frac{2}{4}+\frac{1}{10}\right)+\frac{1}{9}\left[\frac{1}{10}-\frac{1}{16}\right]$$

$$= \frac{1}{18}\cdot\frac{6}{10}+\frac{1}{9}\cdot\underbrace{\frac{16-10}{160}}_{\frac{3}{80}}$$

$$= \underbrace{\frac{1}{30}}+\frac{1}{240}=\frac{9}{240}=\frac{3}{80}=\operatorname{Var}Y$$

Alternative ending:

$$\frac{1}{18}E(1-X)^2=\frac{1}{18}\int_{-\infty}^{\infty}(1-x)^2 f_X(x)\,dx$$

$f_X(x) \to 3(1-x)^2$ for $0<x<1$

$$= \frac{1}{18}\cdot 3\int_0^1(1-x)^4\,dx$$

$$= \frac{1}{6}\int_0^1 u^4\,du=\frac{1}{6}\cdot\frac{1}{5}=\frac{1}{30}$$

Example (a hierarchical model)

Sample a lightbulb at random.

Let X = lifetime of bulb (in years),

Q = quality of bulb (not directly observable).

Suppose that

$$X \mid Q \sim \text{exponential}(Q)$$

$$Q \sim \text{Uniform}(0,2)$$

(Lightbulbs of quality Q have a mean lifetime of Q years.)

The hierarchical statement above is equivalent to saying that

$$f_{X|Q}(x|q) = \frac{1}{q} e^{-x/q} \quad \text{for } x>0, \text{ and}$$

$$f_Q(q) = \frac{1}{2} \quad \text{for } 0<q<2$$

which implies

$$f_{X,Q}(x,q) = f_Q(q) \, f_{X|Q}(x|q)$$

$$= \frac{1}{2q} e^{-x/q} \quad \text{for } x>0,\ 0<q<2.$$

What are the mean and variance of X?

Preliminaries :

From the appendix we find

$\text{Exp}(\beta)$ has mean $= \beta$, variance $= \beta^2$.

$\text{Uniform}(a,b)$ has mean $= \dfrac{a+b}{2}$,

variance $= \dfrac{(b-a)^2}{12}$.

Since $X|Q \sim \text{Exp}(Q)$,

$E(X|Q) = Q$ and $\text{Var}(X|Q) = Q^2$.

Since $Q \sim \text{Uniform}(0,2)$, we know

$EQ = 1$, $\text{Var } Q = \dfrac{2^2}{12} = \dfrac{1}{3}$.

Thus

$EX = E[E(X|Q)] = EQ = 1$, and

$$\begin{aligned}
\text{Var } X &= E[\text{Var}(X|Q)] + \text{Var}(E(X|Q)) \\
&= EQ^2 + \text{Var } Q \\
&= (\text{Var } Q + (EQ)^2) + \frac{1}{3} \\
&= \frac{1}{3} + 1^2 + \frac{1}{3} = \frac{5}{3}
\end{aligned}$$

Example (a hierarchical model)

Sample a leaf at random.
Let X = Area of leaf,
Y = # of bugs on leaf.

Suppose that

$$X \sim \text{Gamma}(\alpha, \beta)$$

$$Y|X \sim \text{Poisson}(\theta X) \qquad (*)$$

where α, β, θ are known values.

(Comment: $(*)$ says the average number of bugs on a leaf is proportional to the area of the leaf.)

---

① What is the (marginal) distn. of Y?

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x)$$

$$= \Bigg(\underbrace{\frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}} I_{(0,\infty)}(x)}_{\text{pdf of Gamma}(\alpha,\beta)} \Bigg) \Bigg(\underbrace{\frac{(\theta x)^y e^{-\theta x}}{y!} I_{\{0,1,2,\ldots\}}(y)}_{\substack{\text{pmf of Poisson} \\ \text{with } \lambda = \theta x}} \Bigg)$$

(Note that X is continuous and Y is discrete.)

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

$$= \frac{\theta^y}{\Gamma(\alpha)\beta^\alpha y!} \underbrace{\int_0^\infty x^{\alpha+y-1} e^{-(\theta+\frac{1}{\beta})x}\,dx}$$

$$= \frac{\Gamma(\alpha+y)}{(\theta+\frac{1}{\beta})^{\alpha+y}}$$ either

by recognizing this as the kernel of a gamma density or by making the substitution $u = (\theta + \frac{1}{\beta})x$ and using the definition of the gamma function.

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot \frac{\Gamma(\alpha+y)}{y!} \cdot \frac{\theta^y}{(\theta+\frac{1}{\beta})^{\alpha+y}} \quad \text{for } y = 0, 1, 2, \ldots$$

This is called a <u>Gamma mixture of Poisson distributions</u>.

② What are the mean and variance of Y?

From the appendix:

Gamma(α, β) has mean $= \alpha\beta$,

variance $= \alpha\beta^2$.

Poisson(λ) has mean $=$ variance $= \lambda$.

Since $Y|X \sim \text{Poisson}(\theta X)$, we know

$$E(Y|X) = \text{Var}(Y|X) = \theta X.$$

Since $X \sim \text{Gamma}(\alpha, \beta)$, we know

$$EX = \alpha\beta, \quad \text{Var}(X) = \alpha\beta^2.$$

Thus

$$EY = E[E(Y|X)] = E(\theta X) = \theta(EX)$$
$$= \theta\alpha\beta,$$

$$\text{Var}\,Y = E(\text{Var}(Y|X)) + \text{Var}(E(Y|X))$$
$$= E(\theta X) + \text{Var}(\theta X)$$
$$= \theta(EX) + \theta^2(\text{Var}X)$$
$$= \theta\alpha\beta + \theta^2\alpha\beta^2.$$

These can be computed from the pmf $f_Y(y)$ found in ①, but this requires a lot more work!

Example (a hierarchical model)

A type of flower comes in two varieties, red and blue.

A packet of seeds contains R seeds for the red variety, and B for the blue.

K seeds are selected at random and planted.

Each seed has probability p of growing and (eventually) producing flowers.

Let X be the number of plants that produce red flowers.

Find the mean and variance of X.

Let Y = # of "red" seeds drawn from the packet.

$$Y \sim \text{Hypergeometric}(N = R+B, M = R, K = k)$$

$$X \mid Y \sim \text{Binomial}(Y, p)$$

From the appendix we obtain

Hypergeometric(N, M, K) has

mean $= KM/N$,

variance $= \dfrac{KM}{N} \cdot \dfrac{(N-M)(N-K)}{N(N-1)}$.

Binomial(n, p) has

mean $= np$, variance $= np(1-p)$

so that

$$EY = \frac{KR}{R+B}, \quad \text{Var}\, Y = \frac{KRB(R+B-K)}{(R+B)^2(R+B-1)}$$

and $E(X|Y) = Yp$, $\text{Var}(X|Y) = Yp(1-p)$.

Thus

$$EX = E[E(X|Y)] = E(Yp) = p(EY)$$

$$= p\left(\frac{KR}{R+B}\right),$$

$$\text{Var}\, X = E(\text{Var}(X|Y)) + \text{Var}(E(X|Y))$$

$$= E(Yp(1-p)) + \text{Var}(Yp)$$

$$= p(1-p)EY + p^2 \text{Var}\, Y$$

$$= p(1-p)\frac{KR}{R+B} + p^2 \frac{KRB(R+B-K)}{(R+B)^2(R+B-1)}$$