$(X, Y) \leftarrow$ 2d r.v.

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

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

$$= \int h(t,y) \frac{f(t,y)}{f_1(t)} \, dy$$

$$f_{Y|X=t}(y)$$

$$\boxed{E[h(X,Y) \mid X] = \xi_h(X)}$$

definition

Results:

$$\mathbb{E}[\mathbb{E}[h(X,Y) \mid X]] = \mathbb{E}[h(X,Y)]$$

double integral

$$= \mathbb{E}[\xi_h(X)] =$$

as an integral w.r.t density of X

$$\mathbb{E}[w(X) h(X,Y) \mid X] = w(X) \, E[h(X,Y) \mid X]$$

Can pull the term just depending on X OUTSIDE conditional expectation

Conditional variance function.

$\underline{Var[h(X,Y)|X]}$ $\rightarrow$ we're going to define this.

$$Var[Z] = \mathbb{E}[(Z - \mathbb{E}Z)^2] = \mathbb{E}Z^2 - (\mathbb{E}Z)^2$$

starting from this expression, define:

$$Var[h(X,Y)|X] = \underline{\mathbb{E}[h^2(X,Y)|X]} - [\mathbb{E}(h(X,Y)|X)]^2$$

Recall that:

$$\mathbb{E}[h^2(X,Y)|X=t] = \int h^2(t,y)\frac{f(t,y)}{f_1(t)}dy$$

etc...

Note:

$$Var[h(X,Y)|X]$$

$$= \mathbb{E}[\underline{(h(X,Y) - \mathbb{E}(h(X,Y)|X))^2}|X]$$

$$= \mathbb{E}[(h(X,Y) - \xi_h(X))^2|X]$$

$$\mathbb{E}\left[\left(h(X,Y)-\xi_h(X)\right)^2 \middle| X\right]$$

$$= \mathbb{E}\left[\left\{h^2(X,Y) - 2\xi_h(X)h(X,Y) + \xi_h(X)^2\right\} \middle| X\right]$$

We can take conditional expectations term by term, and hence:

$$= \mathbb{E}\left[h^2(X,Y)\middle|X\right] - 2\,\underline{\mathbb{E}\left[\xi_h(X)h(X,Y)\middle|X\right]} + \mathbb{E}\left[\xi_h(X)^2\middle|X\right]$$

$$= \mathbb{E}\left[h^2(X,Y)\middle|X\right] - 2\xi_h(X)\,\mathbb{E}\left[h(X,Y)\middle|X\right] + \xi_h(X)^2$$

$$\mathbb{E}\left[h(X,Y)\middle|X\right] = \xi_h(X).$$

giving us:

$$\mathbb{E}\left[h^2(X,Y)\middle|X\right] - 2\xi_h^2(X) + \xi_h^2(X)$$

$$= \mathbb{E}\left[h^2(X,Y)\middle|X\right] - \left(\mathbb{E}\left[h(X,Y)\middle|X\right]\right)^2$$

Variance decomposition formula:

Before I do this, a digression:

(X, Y). random vector. X: d-dimensional
Y: 1 dimensional

$$f(x, y) = f((x_1, \dots, x_d), y)$$

$$x = (x_1, x_2, \dots, x_d).$$

$f_1(x)$: marginal density of x.

$$= \int f(x, y)\, dy$$

function of d variables.

$$\mathbb{E}[h(X, Y) \mid \underline{X}] = \int_{\mathbb{R}^d} h(t, y) \frac{f(t, y)}{f_1(t)}\, dt$$

$$t = (t_1, \dots, t_d)$$

$$\boxed{dt = (dt_1\, dt_2 \dots dt_d)}$$

Variance decomposition formula:

$$\underline{\mathrm{Var}[h(x,y)]} = \mathbb{E}[\mathrm{Var}(h(x,y)|x)] + \mathrm{Var}[\mathbb{E}(h(x,y)|x)]$$

$$\mathrm{Var}[V] = \mathrm{Var}[\mathbb{E}(V|u)] + \mathbb{E}[\mathrm{Var}(V|u)]$$

$V \to \underline{h(x,y)}$, $u \to x$.

for any pair of random variables (u,v) with a joint distribution.

$$\mathrm{Var}[V] = \mathbb{E}[(V - \mathbb{E}V)^2]$$

$$= \mathbb{E}[(V - \mathbb{E}(V|u) + \mathbb{E}(V|u) - \mathbb{E}V)^2]$$

$$= \underline{\mathbb{E}[(V - \mathbb{E}(V|u))^2]} \quad \mathrm{I}$$

$$+ \underline{\mathbb{E}[(\mathbb{E}(V|u) - \mathbb{E}V)^2]} \quad \mathrm{II}$$

$$+ 2\,\mathbb{E}[(V - \mathbb{E}(V|u))(\mathbb{E}(V|u) - \mathbb{E}V)] \quad \mathrm{III}$$

$\mathrm{III} = 0$.

$$\text{I} = \mathbb{E}\left[\left(\underline{V - \mathbb{E}(V|u)}\right)^2\right]$$

$$= \mathbb{E}\left[\underbrace{\mathbb{E}\left[\left(V - \mathbb{E}(V|u)\right)^2 \middle| u\right]}_{\underline{\mathrm{Var}[V|u]}}\right]$$

$$= \mathbb{E}\left[\mathrm{Var}(V|u)\right].$$

$$\text{II} = \mathbb{E}\left[\left(\mathbb{E}(V|u) - \mathbb{E}V\right)^2\right]$$

$$\underline{\mathbb{E}V = \mathbb{E}\left[\mathbb{E}(V|u)\right].}$$

$$\Rightarrow \underline{\text{II}} = \mathbb{E}\left[\left(\underbrace{\mathbb{E}(V|u)}_{\mathcal{K}(u)} - \mathbb{E}\left(\mathbb{E}(V|u)\right)\right)^2\right]$$

$$= \mathbb{E}\left[\mathcal{K}(u) - \mathbb{E}(\mathcal{K}(u))\right]^2$$

$$= \mathrm{Var}\left[\mathcal{K}(u)\right]$$

$$= \mathrm{Var}\left(\mathbb{E}(V|u)\right)$$

$$\frac{\text{III}}{2} = \mathbb{E}\Big[\big(V - \mathbb{E}(V|U)\big)\underbrace{\big(\mathbb{E}(V|U) - \mathbb{E}V\big)}_{\gamma(U)}\Big]$$

$$= \mathbb{E}\Big[\underline{\big(V - \mathbb{E}(V|U)\big)\gamma(U)}\Big]$$

$$= \mathbb{E}\Big[\mathbb{E}\big(\big(V - \mathbb{E}(V|U)\big)\gamma(U)\,\big|\,U\big)\Big]$$

$$= \mathbb{E}\Big[\gamma(U)\,\mathbb{E}\big[\big(V - \mathbb{E}(V|U)\big)\,\big|\,U\big]\Big]$$

$$= \mathbb{E}\Big[\gamma(U)\big[\mathbb{E}(V|U) - \mathbb{E}(V|U)\big]\Big]$$

$$= 0.$$

$$\mathbb{E}\big[\mathbb{E}(V|U)\,\big|\,U\big] = \mathbb{E}(V|U)$$

as $\mathbb{E}(V|U)$ is a function of U.

Exercise

List of n symbols.
Computer searches list sequentially to determine if pre-specified object is present in list, and stops as soon as object is discovered. Otherwise whole list is searched.
Object is in the list with prob p.
What is the expected number of items the program searches to find object?

B : $\mathbb{1}$ (object present in list)

N : # of item searched by computer.

Want to find $\boxed{\mathbb{E}N}$

$N \mid B = 0$ is degenerate at n. w.p. 1.

We'll assume that $N \mid B = 1$ is uniformly distributed over $\{1, 2, \ldots, n\}$

So: $\mathbb{E}N = \underline{\mathbb{E}[\mathbb{E}(N \mid B)]}$

$$= \underset{p}{P(B=1)}\, \mathbb{E}[N \mid B=1] + \underset{1-p}{\underline{P(B=0)}}\, \underset{n}{\underline{\mathbb{E}[N \mid B=0]}}$$

$$= \frac{1}{n} \cdot 1 + \frac{1}{n} \cdot 2 + \cdots + \frac{1}{n} \cdot n$$

$$= \frac{1}{\cancel{n}} \frac{\cancel{n}(n+1)}{2} = \frac{n+1}{2}.$$

<u>So</u>: $\mathbb{E}N = p \cdot \frac{n+1}{2} + (1-p)n$

$$= n - \frac{np}{2} + \frac{p}{2}.$$

<u>Exercise</u>: <u>Fair coin</u>.

<u>Stage 1</u>: Tossed n times. H, the # of Heads is recorded.

Stage 2: Coin tossed again H times. N, the # of Heads is recorded

Total number of heads combining the two stages is $H+N$.

We'd like to find $\mathbb{E}(H+N)$, $\text{Var}(H+N)$

$$\left.\begin{aligned} H &\sim \text{Bin}\left(n, \tfrac{1}{2}\right) \\ N \mid H &\sim \text{Bin}\left(H, \tfrac{1}{2}\right) \end{aligned}\right\}$$

Solve by conditioning on H.

alternative: find joint of (H, N) and then compute the quantities.

try this

$$\mathbb{E}(H+N) = \mathbb{E}H + \mathbb{E}N$$

$$= \frac{n}{2} + \mathbb{E}N$$

$$\mathbb{E}N = \mathbb{E}\left[\mathbb{E}(N \mid H)\right] = \mathbb{E}(H/2)$$

$$= \frac{1}{2} \cdot \mathbb{E}H$$

$$= \frac{n}{4}.$$

So: $\mathbb{E}(N+H) = \dfrac{3n}{4}$

Variance:

$$\mathrm{Var}(N+H) = \mathbb{E}\left[\mathrm{Var}(N+H \mid H)\right] + \mathrm{Var}\left[\mathbb{E}(N+H \mid H)\right]$$

Consider: $\mathbb{E}\left[\mathrm{Var}(N+H \mid H)\right]$

$$= \mathbb{E}\left[\mathrm{Var}(N \mid H)\right]$$

$$= \mathbb{E}\left[H/4\right] = \frac{n}{8}$$

$$\mathrm{Var}\left[\mathbb{E}(N+H \mid H)\right]$$

$$= \mathrm{Var}\left[\mathbb{E}(N \mid H) + H\right]$$

$$= \mathrm{Var}\left[\frac{H}{2} + H\right] = \frac{9}{4} \cdot \mathrm{Var}(H)$$

$$= \frac{9}{4} \times \frac{n}{4} = \frac{9}{16} n.$$

so: $\mathrm{Var}(N+H) = \dfrac{9n}{16} + \dfrac{n}{8} = \dfrac{11n}{16}$

Mixed joint-distribution arising via multistage random experiments.

(Bayesian Inference)

Example: Generate a discrete R.V at Stage 1.

Depending on its value generate continuous random variable at stage 2.

$X \sim f(x)$. p.m.f

X assumes value in countable set $\mathcal{X}$.

So $\sum_{x \in \mathcal{X}} f(x) = 1$

$Y | X = x \sim g_x(y)$ ← p.d.f

$$\boxed{f(x,y)} = f(x) \cdot g_x(y).$$

can't think of this guy as a pure pdf or a pure p.m.f

The marginal density of Y:

$$f_Y(y) = \sum_{x \in \mathcal{X}} f(x) g_x(y).$$

proper p.d.f.

Conditional of X given $Y = y$:

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

p.m.f

Example: $X \sim$ continuous distribution

$Y|X=x \sim$ discrete.

Suppose $X \sim \text{Exp}(\lambda)$.

$Y|X=x \sim \text{Poi}(x)$.

(X, Y)

$$f_{X,Y}(x,y) = \lambda e^{-\lambda x} \mathbb{1}\{x > 0\} \cdot \frac{e^{-x} x^y}{y!} \mathbb{1}\{y \in \mathbb{Z}_{\geq 0}^{+}\}$$

$$f_{X,Y}(x,y) = \frac{\lambda e^{-x(\lambda+1)} x^{y}}{y} \times 1\{x>0, y \in \mathbb{Z}^+\}$$

Marginal of Y.
Conditional of $X \mid Y = y$

$$f_{X|Y=y}(x) = \frac{\dfrac{\lambda e^{-x(\lambda+1)} x^{y}}{\cancel{y!}} 1\{x>0\}}{\underbrace{\left(\int_{[0,\infty)} \dfrac{\lambda e^{-x(\lambda+1)} x^{y}}{\cancel{y!}} \, dx \right)}_{= f_Y(y)}}$$

$f_Y(y) \rightarrow$ is a valid p.m.f

$X \mid Y = y$
$\sim \text{Gamma}(y+1, \lambda+1)$

so:

$$f_{X|Y=y}(x) = \frac{1}{\psi(\lambda,y)} e^{-x(\lambda+1)} x^{y+1-1} \, 1\{x>0\}$$

non-constant of a Gamma$(y+1, \lambda+1)$

$$f_Y(y) = \int_0^\infty \frac{\lambda e^{-x(\lambda+1)} x^y}{y!} \, dx$$

$$= \frac{\lambda}{y!} \int_0^\infty e^{-x(\lambda+1)} x^{y+1-1} \, dx$$

$$= \frac{\lambda}{\cancel{y!}} \frac{\cancel{\Gamma(y+1)}}{(\lambda+1)^{y+1}} = \frac{\lambda}{(\lambda+1)^{y+1}}, \quad y \geq 0.$$

Check that this is a valid p.m.f.