Lecture 21 - Iterated Expectation and BivNormal

Tuesday, November 23, 2021 1:59 PM

Bayes' Theorem: $f(y|x) = \dfrac{f(x|y)f(y)}{f(x)}$

Law of Total Prob:

$$\begin{cases} \text{(discrete)} & f(y) = \sum_x f(y|x)f(x) \\ \text{(cts)} & f(y) = \int f(y|x)f(x)\,dx \end{cases}$$

Ex. $X \sim Exp(\lambda)$

$Y | X = x \sim Pois(x)$ $\quad \dfrac{x^y e^{-x}}{y!}$

What is the dist of Y?

Law of Total Prob.

$$f(y) = \int f(y|x)f(x)\,dx$$

$$= \int_0^\infty \underbrace{\frac{x^y e^{-x}}{y!}}_{Pois(x)} \underbrace{\lambda e^{-\lambda x}}_{Exp(\lambda)}\,dx$$

$a - 1 = y \iff \boxed{a = y+1}$

$\boxed{b = \lambda + 1}$

$$\lambda \int^\infty x^{y} e^{-(\lambda+1)x}\,dx$$

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

$b = \lambda + 1$

looks like Gamma PDF

$$= \frac{\lambda}{y!} \frac{\Gamma(y+1)}{(\lambda+1)^{y+1}} \underbrace{\int_0^\infty \frac{x^{y} e^{-(\lambda+1)x} (\lambda+1)^{y+1}}{\Gamma(y+1)} \, dx}_{\text{Integrate to 1}}$$

$$\frac{(bx)^{a-1} b e^{-bx}}{\Gamma(a)}$$

$$= \frac{\lambda \cancel{\Gamma(y+1)}}{\cancel{y!}(\lambda+1)^{y+1}}$$

$\Gamma(y+1) = y!$

$$= \boxed{\frac{\lambda}{(\lambda+1)^{y+1}} \quad \text{for } y = 0, 1, 2, \ldots}$$

Ex, $Y \sim \text{Pois}(\lambda)$

$X \mid Y = y \sim \text{Bin}(y, p)$ $\quad 0 < p < 1$

Note: $0 \leq X \leq Y$

What is the dist of X?

$$f(x) = \sum_y f(x|y) f(y) = \sum_{y=x}^{\infty} \binom{y}{x} p^x (1-p)^{y-x} \frac{\lambda^y e^{-\lambda}}{y!}$$

$$\binom{y}{x} \frac{1}{y!} = \frac{\cancel{y!}}{x!(y-x)!} \frac{1}{\cancel{y!}}$$

$$= \frac{\lambda^x p^x e^{-\lambda}}{x!} \sum_{y=x}^{\infty} \frac{1}{(y-x)!} (1-p)^{y-x} \lambda^{y-x}$$

$$\frac{\lambda^x p^x e^{-\lambda}}{x!} \sum_{y=x} \frac{\quad}{(y-x)!}$$

$$= \frac{\lambda^x p^x e^{-\lambda}}{x!} \underbrace{\sum_{y=0}^{\infty} \frac{1}{y!}\left[(1-p)\lambda\right]^y}_{e^{(1-p)\lambda}}$$

$$\boxed{e^z = \sum_{i=0}^{\infty} \frac{z^i}{i!}}$$

$$= \frac{(\lambda p)^x \cancel{e^{-\lambda}}}{x!} e^{(\cancel{1}-p)\lambda} = \boxed{\frac{(\lambda p)^x e^{-\lambda p}}{x!}} = f(x)$$

Pois (λp)

Theorem: Iterated Expectation

If X and Y are RVs then

fn of Y

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[X|Y]\right].$$

$$\mathbb{E}[X|Y=y] = \int x f(x|y)\,dx \in \mathbb{R}$$

For each $y \in \mathbb{R}$ we get some value

This defines a fn

$$g(y) = \mathbb{E}[X|Y=y]$$

We can plug Y into g to get $g(Y)$.

Might denote: $\quad g(Y) = \dots ?$

Might denote

$\cancel{\mathbb{E}[X|Y=Y]}$ awkward

denote $\mathbb{E}[X|Y]$.

Ex, $g(y)=y^2$

$g(Y)=Y^2$

Basically:

1. $E[X|Y=y]$ is a number

2. $E[X|Y]$ is a RV

Ex, $\mathbb{E}[X|Y=y]=y^2$

then $E[X|Y]=Y^2$.

Ex, $Y \sim \text{Pois}(\lambda)$

$X|Y=y \sim \text{Bin}(y,p)$

$\boxed{\mathbb{E}X = E[E[X|Y]]}$

what is $E[X]$?

1. get $\mathbb{E}[X|Y=y]=yp$

2. get $E[X|Y]=Yp$

3. $\mathbb{E}\,\mathbb{E}[X|Y] = \mathbb{E}[Yp] = p\,\mathbb{E}[Y] = p\lambda$

③ $\mathbb{E}\,\mathbb{E}[X|Y] = \mathbb{E}[Y p] = p\,\mathbb{E}[Y] = \underline{p\lambda}$

Ex. $P \sim Beta(\alpha, \beta)$

$X|P=p \sim Bin(n,p)$

$\mathbb{E}X$?

① $\mathbb{E}[X|P=p] = np$

② $\mathbb{E}[X|P] = nP$

③ $\mathbb{E}X = \mathbb{E}[\mathbb{E}[X|P]] = \mathbb{E}[nP] = n\mathbb{E}P$

$\boxed{= n\frac{\alpha}{\alpha+\beta}}$

Pf. (iterated expectation) — cts case

① $f(x) = \int f(x,y)\,dy$

② $f(x|y) = \frac{f(x,y)}{f(y)} \iff \underline{f(x,y) = f(x|y)f(y)}$

③ $\mathbb{E}[X|Y=y] = \int x f(x|y)\,dx$

$\mathbb{E}[X] = \int x f(x)\,dx \overset{①}{=} \int x \int f(x,y)\,dy\,dx$

$\overset{②}{=} \int x \int f(x|y) f(y)\,dy\,dx$

$$= \int \underbrace{\int x f(x|y)\,dx}_{\mathbb{E}[X|Y=y] = g(y)} f(y)\,dy \quad \text{(rearrange)}$$

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

$$= \mathbb{E}[g(Y)]$$

$$= \mathbb{E}[\mathbb{E}[X|Y]]$$

notation

Theorem: Law of Total Variance

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathbb{E}[X|Y])$$

simular: Calc $\mathrm{Var}(X|Y=y)$ promote y to Y.

Ex. $P \sim \mathrm{Beta}(\alpha, \beta)$

$X|P=p \sim \mathrm{Bin}(n, p)$

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X|P)] + \mathrm{Var}(\mathbb{E}[X|P])$$

① $\mathbb{E}[X|P=p] = np$

$\mathrm{Var}(X|P=p) = np(1-p)$

① $E[X|P=p] = np$

$Var(X|P=p) = np(1-p)$

② $E[X|P] = nP$

$Var(X|P) = nP(1-P)$

$E[nP(1-P)] + Var(nP)$

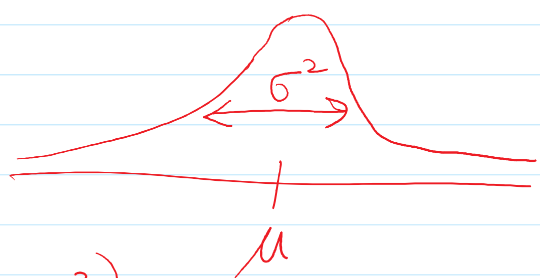
$= n(\underline{E[P]} - \underline{E[P^2]}) + n^2\,\underline{Var(P)}$

$\cdots$

$$= n\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} + \frac{n^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Bivariate Normal

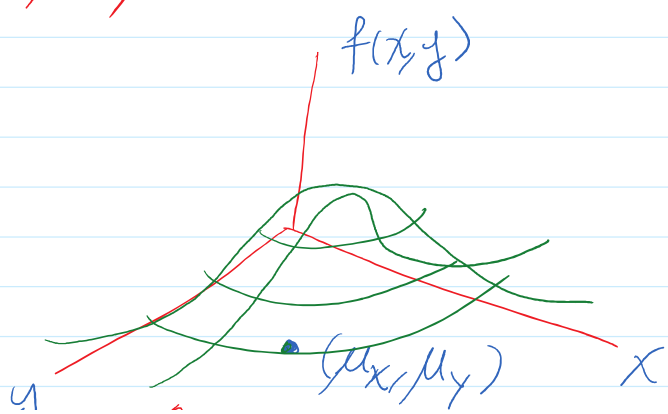
Univariate: $N(\mu, \sigma^2)$

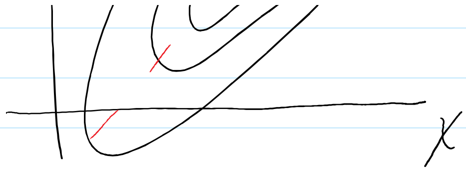


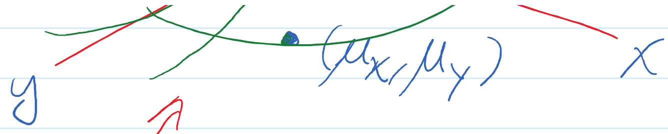
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

Bivariate:

$\rho \gg 0$



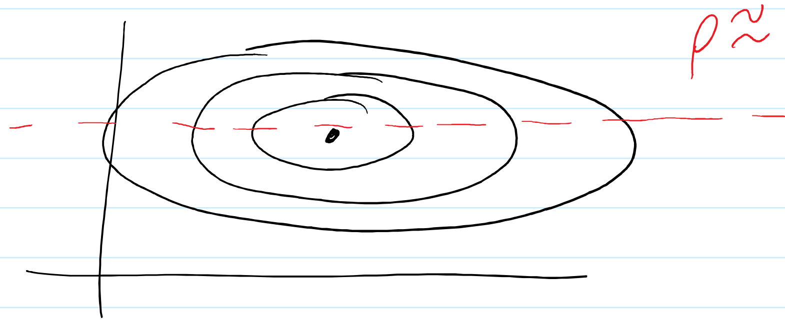




$\sigma_x^2, \sigma_y^2, \rho$

cor btwn X and Y

$\rho \ll 0$



$\rho \approx 0$

<u>PDF:</u> $(X, Y) \sim \text{BivN}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$

$$f(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2} \frac{1}{\sqrt{1-\rho^2}} \left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) \right] \right\}$$

$\mu = (\mu_x, \mu_y)$ mean vecter

covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} Var(X) & Cov(X, Y) \\ Cov(X, Y) & Var(Y) \end{bmatrix}$$

$z = (x, y)$

PDF:

$$f(z) = \frac{1}{2\pi} \frac{1}{\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} (z-\mu)^T \Sigma^{-1} (z-\mu) \right)$$

[Uni: $f(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma^2}} \exp\left(-\frac{1}{2} (x-\mu)(\sigma^2)^{-1}(x-\mu) \right)$]

Facts:

1. $X \sim N(\mu_X, \sigma_X^2)$
   $Y \sim N(\mu_Y, \sigma_Y^2)$

2. $\mathrm{Cor}(X, Y) = \rho$

3. $aX + bY \sim N(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_X\sigma_Y\rho)$

4. $(X, Y) \sim \mathrm{BivN} \iff \forall a, b,\ aX + bY \sim N$

5. Prev: If $X \perp\!\!\!\perp Y$ then $\mathrm{Cor}(X, Y) = 0$

if $(X, Y) \sim \mathrm{BivN}$ and $\rho = 0$ then $X \perp\!\!\!\perp Y$.