$E(Y|X)$, $\mathrm{cov}(X,Y)$, independence

$$\mathrm{cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$
$$= E(XY) - E(X)E(Y)$$

$$\rho = \mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{SD(X)\,SD(Y)}$$

$\rho^2 \le 1$ $\leftarrow$ ρ is the correlation

$\underbrace{\mathrm{cov}(X,Y)=0}_{E(XY)=E(X)E(Y)}$ then X & Y are uncorrelated

Recall X & Y are independent iff

$$E[g(X)\,h(Y)] = E[g(X)]\,E[h(Y)],$$
"$\forall$" g, h

Properties of cov

- $\mathrm{cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$
- $\mathrm{cov}(aX, bY) = ab\, \mathrm{cov}(X, Y)$
- $\mathrm{cov}(X+c, Y+d) = \mathrm{cov}(X, Y)$
- $\mathrm{cov}\left(\sum_{i=1}^{m} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i,j} \mathrm{cov}(X_i, Y_j)$

$$\underset{\sim}{X}, \quad \underbrace{E(\underset{\sim}{X})}_{\underset{\sim}{\mu}} = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_m) \end{pmatrix}$$

(same for a random matrix)

$$\mathrm{Var}(\underset{\sim}{X}) = E\Big[\underbrace{(\underset{\sim}{X} - \underset{\sim}{\mu})}_{m \times 1}\underbrace{(\underset{\sim}{X} - \underset{\sim}{\mu})'}_{1 \times m}\Big] = E(\underset{\sim}{X}\underset{\sim}{X}') - \underset{\sim}{\mu}\underset{\sim}{\mu}'$$

$$\|$$

$$\Sigma$$

$m \times m$

$$= \{\mathrm{cov}(X_i, X_j)\}_{i,j=1}^{m}$$

<u>variance covariance matrix</u>

Note $\Sigma = \Sigma'$

diagonal terms = variance of components

Recall Lemma - $E(aX+b) = a\,E(X) + b$

- $Var(aX+b) = a^2\,Var(X)$

In the vector case

Lemma - $E(\overset{\text{matrix}}{A}\,\underset{\sim}{X} + \underset{\sim}{b}) = A\,E(\underset{\sim}{X}) + \underset{\sim}{b}$

- $Var(A\underset{\sim}{X} + \underset{\sim}{b}) = A\,Var(\underset{\sim}{X})\,A'$

Note $(AB)' = B'A'$

$\det(A) = \det(A')$

Application

$$Var(X_1 + \cdots + X_n) = Var(\overbrace{\underset{\sim}{1}'\underset{\sim}{X}}^{1\cdot X}) = \underset{\sim}{1}'\,Var(\underset{\sim}{X})\,\underset{\sim}{1}$$

$$= \text{sum of all elements of } \Sigma$$

$$= \sum_{i,j} cov(X_i, X_j)$$

<u>Notice</u> $X_1, X_2, \dots, X_n$ uncorrelated

$\Rightarrow \operatorname{cov}(X_i, X_j) = 0$ if $i \neq j$

+ then

$$\operatorname{Var}(X_1 + \dots + X_n) = \operatorname{Var}(X_1) + \dots + \operatorname{Var}(X_n)$$

eg $U \sim \text{Poisson}(\lambda_1)$, $V \sim \text{Poisson}(\lambda_2)$, $W \sim \text{Poisson}(\lambda_3)$

$$X = U + V$$

$$Y = V + W$$

Assume U, V + W are independent.

$$\begin{aligned} \operatorname{cov}(X, Y) &= \operatorname{cov}(U + V, V + W) \\ &= \operatorname{cov}(U, V) + \operatorname{cov}(U, W) \\ &\quad + \operatorname{cov}(V, V) + \operatorname{cov}(V, W) \\ &= \operatorname{cov}(V, V) = \lambda_2 \end{aligned}$$

<u>or</u>

$$E(XY) = \sum_{x,y} xy\, f(x, y) \; ; \quad \begin{aligned} E(X) &= \lambda_1 + \lambda_2 \\ E(Y) &= \lambda_2 + \lambda_3 \end{aligned}$$

weeks later still working
years " " "

or

joint ~~pgf~~ $G(s_1, s_2) = E(s_1^X s_2^Y)$

(joint mgf $m(t_1, t_2) = E(e^{t_1 X} e^{t_2 Y})$)

$$= E\left(s_1^{U+V} s_2^{V+W}\right)$$

$$= E\left(s_1^{U} (s_1 s_2)^{V} s_2^{W}\right)$$

$$= E(s_1^U)\, E\left[(s_1 s_2)^V\right] E(s_2^W)$$

$$= e^{\lambda_1(s_1-1)} e^{\lambda_2(s_1 s_2-1)} e^{\lambda_3(s_2-1)}$$

E(XY)?

$$G(s_1, s_2) = E(s_1^X s_2^Y)$$

$$\frac{\partial^2 G(s_1, s_2)}{\partial s_2 \partial s_1} = E\left(X s_1^{X-1} Y s_2^{Y-1}\right)$$

Set $s_1 = s_2 = 1$ to get

$$\left.\frac{\partial^2 G}{\partial s_2 \partial s_1}\right|_{s_1 = s_2 = 1} = E(XY)$$

& this isn't too hard. Verify that you can get $\text{cov}(X, Y)$ like this.

$E(Y|X)$

$$f(y|x) \longrightarrow \underbrace{E(Y \mid X = x)}_{r(x)}$$

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

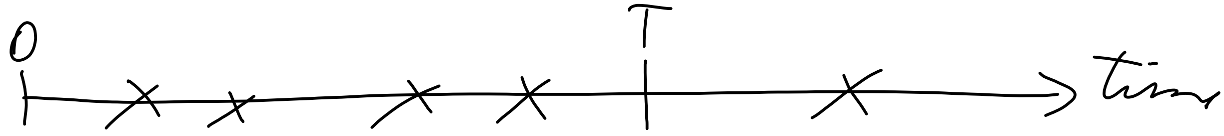
$$\sum_y y\, f(y|x)$$

$r(X)$ is $E(Y|X)$

It's easy to show $E[r(X)] = E(Y)$

$$E[r(X)] = \int_{-\infty}^{\infty} r(x) f(x)\, dx$$

eg



Assume # of x's in an interval of length l is Poisson(λl) & that #'s x in intervals which don't overlap are independent.

rate

- point process
- Poisson point process
- marked point process

Assume damages $D_1, D_2, \ldots$ iid with mean μ & these are independent of the # of x's.

Let N = # of x's from 0 to T
$\sim$ Poisson(λT)

$$\underbrace{\text{Total damages}}_{S_N} = \sum_{k=1}^{N} D_k \qquad (S_0 = 0)$$

$$E(S_N) \neq \sum_{k=1}^{N} E(D_k) = N\mu$$

$$E(S_N \mid N) = \sum_{k=1}^{N} E(D_k) = N\mu$$

$$E(S_N) = E\left[E(S_N \mid N)\right] = \underbrace{E(N)}_{\lambda T}\,\mu$$

For $Var(S_N)$ use formula in the text ~~or~~

$$E(e^{tS_N}) = E\Big[\underbrace{E(e^{tS_N} \mid N)}_{[\underbrace{E(e^{tD_1})}_{m_{D_1}(t)}]^N}\Big]$$

$$= E\left(m_{D_1}(t)^N\right)$$

$$= G_N\left(m_{D_1}(t)\right) = e^{\lambda T[m_{D_1}(t) - 1]}$$

Now get moments . . .

$E(Y|X)$? An approximation / predictor of Y using f'n of X. Denote it by $\hat{Y}$.

$\hat{Y}$ minimizes $E(Y - \text{f'n of } X)^2$

or

Solve
$$E[Y\,h(X)] = E[\hat{Y}\,h(X)], \quad \text{"}\forall\text{" } h$$

$\hat{Y} \to E(Y|X)$

If $h = 1$ then
$$\boxed{E(Y) = E(\hat{Y})}$$

Calculation of $\hat{Y}$?
- get $r(x)$
- $\hat{Y} = r(X)$

eg Back to the bivariate Poisson.

$$X = U + V$$

$$Y = V + W$$

Fix $X = x$ then

$$V \sim \text{binomial}\left(x,\ p = \frac{\lambda_2}{\lambda_1+\lambda_2}\right)$$

& hence

$$\underbrace{E(Y \mid X = x)}_{r(x)} = x \frac{\lambda_2}{\lambda_1+\lambda_2} + \lambda_3$$

So

$$r(x) = (\ ) + (\ )\,x$$

$$\Rightarrow E(Y \mid X) = (\ ) + (\ )\,X$$

vector normal

$$\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \qquad \underset{\sim}{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \leftarrow \text{iid } N(0,1)$$

Let $\underset{\sim}{\mu} = E(\underset{\sim}{Y}) = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$

$$\Sigma = \mathrm{Var}(\underset{\sim}{Y}) = \begin{pmatrix} \mathrm{cov}(Y_1, Y_1) & \mathrm{cov}(Y_1, Y_2) \\ \mathrm{cov}(Y_2, Y_1) & \mathrm{cov}(Y_2, Y_2) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix} \quad - \quad \rho = \frac{\mathrm{cov}(Y_1, Y_2)}{\sigma_1 \sigma_2}$$

$$T = \left(\diagdown \; 0 \right)$$

Find $T \ni$

$$T T' = \Sigma$$

Try

$$T = \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2 \rho & \sigma_2 \sqrt{1-\rho^2} \end{pmatrix}$$

Set

$$\underset{\sim}{Y} = \underset{\sim}{\mu} + T\,\underset{\sim}{Z}$$

$$\Rightarrow E(\underset{\sim}{Y}) = \underset{\sim}{\mu}$$

$$Var(\underset{\sim}{Y}) = T\,\overbrace{Var(\underset{\sim}{Z})}^{\begin{pmatrix}1&0\\0&1\end{pmatrix}}\,T' = TT' = \Sigma$$

$\underset{\sim}{Y}$ is bivariate normal

$$\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}, \Sigma)$$

Notice

$$Y_1 = \mu_1 + \sigma_1 Z_1$$

$$Y_2 = \mu_2 + \sigma_2 \rho Z_1 + \sigma_2 \sqrt{1-\rho^2}\, Z_2$$

$\underline{r(y_1) = E(Y_2 \mid Y_1 = y_1)}$

Sol'n #1 — obtain $f(y_1, y_2)$ via change of variables

— we know $f(y_1)$

— then $f(y_2 | y_1) = \dfrac{f(y_1, y_2)}{f(y_1)}$

— $r(y_1) = \displaystyle\int_{-\infty}^{\infty} y_2 \, f(y_2 | y_1) \, dy_2$

Please do it.

<u>Sol'n # 2</u>

Fix $Y_1 = y_1 \Rightarrow Z_1 = \dfrac{y_1 - \mu_1}{\sigma_1}$

$$\Rightarrow Y_2 = \mu_2 + \sigma_2 \rho \left(\frac{y_1 - \mu_1}{\sigma_1} \right) + \sigma_2 \sqrt{1-\rho^2} \, Z_2$$

$$\Rightarrow E(Y_2 | Y_1 = y_1) = \overbrace{\left(\mu_2 - \rho \mu_1 \frac{\sigma_2}{\sigma_1} \right)}^{a} + \rho \frac{\sigma_2}{\sigma_1} y_1$$

$$+ (\ ) \underbrace{E(Z_2 | Y_1 = y)}_{0 = E(Z_2)}$$

$\because$ Z_2 & Y_1 are ind

$\therefore$ $r(y_1) = a + \rho \dfrac{\sigma_2}{\sigma_1} y_1$

Also
$$E(Y_2 \mid Y_1) = a + \rho \frac{\sigma_2}{\sigma_1} Y_1$$

Notice that

$$\underset{\sim}{c}' \underset{\sim}{Y} = c_1 Y_1 + c_2 Y_2$$ (linear combination of Y_1 & Y_2)

$$\sim N(\,?\,,\,??\,),$$

where
$$? = \underset{\sim}{c}' E(\underset{\sim}{Y}) = \underset{\sim}{c}' \underset{\sim}{\mu} = c_1 \mu_1 + c_2 \mu_2$$

$$?? = \underset{\sim}{c}' \Sigma \underset{\sim}{c}$$

Note $X \sim N(\mu, \sigma^2)$
$E(e^X) = e^{\mu} e^{\sigma^2/2}$

mgf of $\underset{\sim}{Y}$

$$m(\underset{\sim}{t}) = E\left(e^{\underset{\sim}{t}'\underset{\sim}{Y}}\right) = e^{\underset{\sim}{t}'\underset{\sim}{\mu}} e^{\underset{\sim}{t}' \Sigma \underset{\sim}{t}/2}$$

Def'n $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}, \Sigma)$ if

$$m(\underset{\sim}{t}) = e^{\underset{\sim}{t}'\underset{\sim}{\mu}} e^{\underset{\sim}{t}' \Sigma \underset{\sim}{t}/2}$$

$\Leftrightarrow$ $\underset{\sim}{Y} \overset{d}{=} \underset{\sim}{\mu} + T \underset{\sim}{Z}$, where

$$\Sigma = T T'$$

+ the components of $\underset{\sim}{Z}$ are iid $N(0,1)$

The pdf of a $N(\underset{\sim}{\mu}, \Sigma)$ is

$$f(\underset{\sim}{y}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{n} \frac{1}{\sqrt{\det(\Sigma)}} \exp\left[-\frac{(\underset{\sim}{y}-\underset{\sim}{\mu})' \Sigma^{-1} (\underset{\sim}{y}-\underset{\sim}{\mu})}{2}\right]$$

<u>Note</u> — $\underset{m \times m}{A}$, $A^{-1}A = AA^{-1} = I$

— $(AB)^{-1} = B^{-1}A^{-1}$

— $\det(AB) = \det(A)\det(B)$

— $\det(A') = \det(A)$

— $\det(\Sigma) = \det(TT') = \det(T)\det(T') = (\det(T))^2$

$\Rightarrow \det(T) = \sqrt{\det(\Sigma)}$

— $\det(A^{-1}) = \dfrac{1}{\det(A)}$

<u>Proposition</u> If $\underset{\sim}{Y} \sim N(\underset{\sim}{\mu}, \Sigma)$ then

$$A\underset{\sim}{Y} + \underset{\sim}{b} \sim N(A\underset{\sim}{\mu} + \underset{\sim}{b}, A\Sigma A')$$

<u>Proof</u> Use mgf's

<u>multinomial</u>

Toss a 2-sided coin with

$P(\text{side } \#1) = p_1$, $P(\text{side } \#2) = p_2$

Let $\underset{2\times 1}{\underset{\sim}{Z}} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$ $\begin{matrix} \leftarrow I_{\{\text{side}\#1\}} \\ \leftarrow I_{\{\text{side}\#2\}} \end{matrix}$

Only possible values for $\underset{\sim}{Z}$ are

prob $p_2 \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \leftarrow$ prob p_1

pgf of $\underset{\sim}{Z}$ is $G(\underset{\sim}{s}) = E\left(s_1^{Z_1} s_2^{Z_2}\right)$

$= p_1 s_1 + p_2 s_2$

Now let $\underset{\sim}{Z}_1, \cdots, \underset{\sim}{Z}_m$ be iid $\underset{\sim}{Z}$ & set

$$\underset{\sim}{Y} = \underset{\sim}{Z}_1 + \cdots + \underset{\sim}{Z}_m$$

$$\Rightarrow \quad G_{\underset{\sim}{Y}}(\underset{\sim}{s}) = (p_1 s_1 + p_2 s_2)^m$$

Note $Y_1 + Y_2 = m$ &

$$\underbrace{P(Y_1 = y_1, Y_2 = y_2)}_{f(y_1, y_2)} = \binom{m}{y_1, y_2} p_1^{y_1} p_2^{y_2}, \quad y_1 + y_2 = m$$

$$\binom{m}{y_1, y_2} = \frac{m!}{y_1! \, y_2!}$$

<u>k-sided die</u> probabilities $p_1, \cdots, p_k$

$$\Rightarrow \quad G_{\underset{\sim}{Y}}(\underset{\sim}{s}) = (p_1 s_1 + \cdots + p_k s_k)^m$$

$$\Rightarrow \quad f(\underset{\sim}{y}) = \binom{m}{y_1, \cdots, y_k} p_1^{y_1} \cdots p_k^{y_k}, \quad y_1 + \cdots + y_k = m$$

$$\underset{\sim}{Y} \sim \text{multinomial}(m; \underset{\sim}{p}), \quad \underset{\sim}{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_k \end{pmatrix}$$