Exercises - Lecture 18 - Solutions

Ex III.7.1 $P(\exp(tX) \geq k) \leq E(\exp\{tX\})/k$ which as we will show is $m_X(t)/k$ where m_X is the mgf of X.

Ex III.7.2 $P(|X| \geq k) \overset{\text{Markov}}{\leq} E(|X|)/k$ but also $P(|X| \geq k) = P(X^2 \geq k^2) \overset{\text{Markov}}{\leq} E(X^2)/k^2$

When $X \sim$ exponential (1) then $P(X \geq 2)$

$= \int_2^\infty e^{-x}\,dx = -e^{-x}\Big|_2^\infty = e^{-2} = 0.135$

while $E(|X|)/2 = \int_0^\infty x e^{-x}\,dx/2 = 1/2 = 0.500$

and $E(X^2)/4 = \int_0^\infty x^2 e^{-x}\,dx/4 = \Gamma(3)/4 = 2/4 = 0.500$

Ex III.7.3 Suppose $E((Y-a-bX)^2)$

$= \min_{a_0, b_0 \in R} E((Y-a_0-b_0X)^2)$

$= \min_{a_0, b_0 \in R} E((Y-\mu_Y+\mu_Y-a_0-b_0(X-\mu_X)-b_0\mu_X)^2)$

$= \min_{a_0, b_0 \in R} E(((Y-\mu_Y)-(a_0-\mu_Y)+b_0\mu_X)-b_0(X-\mu_X))^2$

and so $a_Y = a_0-\mu_Y+b\mu_X$, $b_Y = b_0$ minimizing

$E(((Y-\mu_Y)-a_{00}-b_{00}(X-\mu_X))^2)$ over

all constants a_{00}, b_{00}.

Ex III.7.4

(i) Based on Ex III.7.3 we can assume $E(X) = E(Y) = 0$ and find a, b minimizing $E((Y-a-bX)^2)$.

Putting $c_{XY} = \sigma_Y \rho_{XY} / \sigma_X$ we have

$$E(Y - c_{XY} X) = E(Y) - c_{XY} E(X) = 0$$

$$\mathrm{COV}(Y - c_{XY}X, a+bX) = E((Y - c_{XY}X)(a+bX))$$

$$= E(aY + bXY - ac_{XY}X - bc_{XY}X^2)$$

$$= aE(Y) + bE(XY) - ac_{XY}E(X) - bc_{XY}E(X^2)$$

and using $E(X) = E(Y) = 0$, $E(X^2) = \sigma_X^2$, $E(XY) = \sigma_{XY}$

$$= b\sigma_{XY} - bc_{XY}\sigma_X^2 = b(\sigma_{XY} - \sigma_X\sigma_Y\rho_{XY}) = 0$$

since $\sigma_X\sigma_Y\rho_{XY} = \sigma_{XY}$.

$$E((Y-a-bX)^2) = E((Y - c_{XY}X - a - (b - c_{XY})X)^2)$$

$$= E((Y - c_{XY}X)^2) - 2E((Y - c_{XY}X)(a + (b - c_{XY})X))$$

$$+ E((a + (b - c_{XY})X)^2)$$

by above results

$$= \mathrm{Var}(Y - c_{XY}X) + E(a^2 + 2a(b - c_X)X + (b - c_{XY})^2X^2)$$

$= Var(Y - c_{xy}X) + a^2 + (b - c_{xy})^2 Var(X)$

and this is minimized over (a,b) by $a=0$ and $b = c_{xy}$. Therefore the best affine predictor of Y from X is $c_{xy}X$.

(ii) In general (without assuming $E(X) = E(Y) = 0$) the best affine predictor of Y from X is the $\mu_Y + c_{xy}(X - \mu_X)$.

(iii) By (i) the proportion of variation in Y explained by the best affine predictor is (taking $a=0$, $b = c_{xy}$)

$Var(Y - c_{xy}X) = E((Y - c_{xy}X)^2)$

$= E(Y^2) - 2c_{xy}E(XY) + c_{xy}^2 E(X^2)$

$= \sigma_Y^2 - 2c_{xy}\sigma_{xy} + c_{xy}^2\sigma_X^2$

$= \sigma_Y^2 - 2\frac{\sigma_Y}{\sigma_X}\rho_{xy}\sigma_{xy} + \sigma_Y^2\rho_{xy}^2$

$= \sigma_Y^2 - 2\sigma_Y^2\rho_{xy}\frac{\sigma_{xy}}{\sigma_X\sigma_Y} + \sigma_Y^2\rho_{xy}^2$

$= \sigma_Y^2(1 - 2\rho_{xy}^2 + \rho_{xy}^2) = \sigma_Y^2(1 - \rho_{xy}^2)$

is the residual variation in Y. So the proportion of the total variation in Y

explained by the best affine predictor is

$$\frac{\sigma_Y^2 - \sigma_Y^2(1-\rho_{XY}^2)}{\sigma_Y^2} = \rho_{XY}^2 .$$

(iv) We determined in an earlier result that

$$E_{Y|X}(Y|X)(x) = \mu_Y + \sigma_{XY}(\sigma_X^2)^{-1}(x-\mu_X)$$
$$= \mu_Y + \frac{\sigma_Y}{\sigma_X}\rho_{XY}(x-\mu_X)$$

which is the best affine predictor of Y from X.