Stats 100A Homework 4

Jiaping Zeng

5/28/2020

Problem 1

(a) $P(X=1) = p(1,1) + p(1,2) + p(1,3) = \boxed{0.4}$
$P(X=2) = p(2,1) + p(2,2) + p(2,3) = \boxed{0.6}$
$P(Y=1) = p(1,1) + p(2,1) = \boxed{0.2}$
$P(Y=2) = p(1,2) + p(2,2) = \boxed{0.3}$
$P(Y=3) = p(1,3) + p(2,3) = \boxed{0.5}$

(b) $P(X=1|Y=1) = \dfrac{p(1,1)}{0.2} = \boxed{\dfrac{1}{2}}$
$P(X=1|Y=2) = \dfrac{p(1,2)}{0.3} = \boxed{\dfrac{1}{3}}$
$P(X=1|Y=3) = \dfrac{p(1,3)}{0.3} = \boxed{\dfrac{2}{5}}$
$P(X=2|Y=1) = \dfrac{p(2,1)}{0.2} = \boxed{\dfrac{1}{2}}$
$P(X=2|Y=2) = \dfrac{p(2,2)}{0.3} = \boxed{\dfrac{2}{3}}$
$P(X=2|Y=3) = \dfrac{p(2,3)}{0.5} = \boxed{\dfrac{3}{5}}$
$P(Y=1|X=1) = \dfrac{p(1,1)}{0.4} = \boxed{\dfrac{1}{4}}$
$P(Y=1|X=2) = \dfrac{p(2,1)}{0.6} = \boxed{\dfrac{1}{6}}$
$P(Y=2|X=1) = \dfrac{p(1,2)}{0.4} = \boxed{\dfrac{1}{4}}$
$P(Y=2|X=2) = \dfrac{p(2,2)}{0.6} = \boxed{\dfrac{1}{3}}$
$P(Y=3|X=1) = \dfrac{p(1,3)}{0.4} = \boxed{\dfrac{1}{2}}$
$P(Y=3|X=2) = \dfrac{p(2,3)}{0.6} = \boxed{\dfrac{1}{2}}$

(c) $E(X) = 1 * 0.4 + 2 * 0.6 = 1.6$
$E(Y) = 1 * 0.2 + 2 * 0.3 + 3 * 0.5 = 2.3$

$E(XY) = 1*1*0.1 + 1*2*0.1 + 1*3*0.2 + 2*1*0.1 + 2*2*0.2 + 2*3*0.3 = 3.7$
$Cov(X,Y) = E(XY) - E(X)E(Y) = 3.7 - 1.6*2.3 = \boxed{0.02}$

Problem 2

(a) $\mu = 0, \sigma^2 = 1 \implies f(x) = \boxed{\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}}$

$\mu = \rho x, \sigma^2 = 1 - \rho^2 \implies f(y|x) = \boxed{\frac{e^{\frac{-(y-\rho x)^2}{2(1-\rho^2)}}}{\sqrt{2\pi(1-\rho^2)}}}$

$f(x,y) = f(x)f(y|x) = \boxed{\frac{e^{\frac{-(y-\rho x)^2-(1-\rho^2)x^2}{2(1-\rho^2)}}}{2\pi\sqrt{(1-\rho^2)}}}$

(b) $E(Y|X=x) = E(\rho x + \epsilon) = E(\rho x) = \boxed{\rho x}$

$Var(Y|X=x) = Var(\rho x + \epsilon) = Var(\epsilon) = \boxed{1-\rho^2}$

(c) $E(Y) = E(\rho X + \epsilon) = \boxed{0}$

$Var(Y) = Var(\rho X + \epsilon) = \rho^2 Var(X) + Var(\epsilon) = \boxed{1}$

$Cov(X,Y) = E(XY) = \rho E(X^2) + E(X\epsilon) = \boxed{\rho}$

Problem 3

(a) $E(X+Y)$
$= \sum_x \sum_y (x+y)p(x,y)$
$= \sum_x x \sum_y p(x,y) + \sum_y y \sum_x p(x,y)$
$= \sum_x xp(x) + \sum_y yp(y)$
$= E(X) + E(Y)$

(b) Let $\mu_{X+Y} = E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y$, then:
$Var(X+Y)$
$= E[(X+Y-\mu_{X+Y})^2]$
$= E[((X-\mu_X)+(Y-\mu_Y))^2]$
$= E[(X-\mu_X)^2] + E[(Y-\mu_Y)^2] + 2E[(X-\mu_X)(Y-\mu_Y)]$
$= Var(X) + Var(Y) + 2Cov(X,Y)$

(c) $Cov(X,Y)$
$= E[(X-\mu_X)(Y-\mu_Y)]$
$= \sum_X (x-\mu_X)p_X(x) \sum_Y (y-\mu_Y)p_Y(y)$
$= (\sum_X xp_X(x) - \mu_X)(\sum_Y yp_Y(y) - \mu_Y)$
$= (E(X) - \mu_X)(E(Y) - \mu_Y)$
$= 0$

The reverse is not always true. Counter example: $X\ Unif[-1,1], Y = X^2$. We can show that

$Cov(X, Y) = 0$, though X and Y are not independent, as follows:

$Cov(X, Y)$

$= E(XY) - E(X)E(Y)$

$= E(X^3) - E(X)E(X^2)$

$= 0$

Problem 4

(a) $E(X_i) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = 3.5$

$\implies E(\bar{X}) = \frac{100 * 3.5}{100} = \boxed{3.5} = \mu$

$Var(X_i) = (1-3.5)^2 * \frac{1}{6} + (2-3.5) * \frac{1}{6} + (3-3.5)^2 * \frac{1}{6} + (4-3.5)^2 * \frac{1}{6} + (5-3.5)^2 * \frac{1}{6} + (6-3.5)^2 * \frac{1}{6} = \frac{17.5}{6}$

$\implies Var(\bar{X}) = \frac{100 * \frac{17.5}{6}}{100^2} = \boxed{\frac{7}{240}} = \sigma^2$

(b) $SD(\bar{X}) = \sigma = \sqrt{\frac{7}{240}} \implies \bar{X} \in [\mu - 2\sigma, \mu + 2\sigma] \implies \boxed{\bar{X} \in [3.158, 3.842]}$

(c) $\sigma = \sqrt{\frac{10000 * \frac{17.5}{6}}{10000^2}} \implies \boxed{\bar{X} \in [3.466, 3.534]}$

(d) $E(X) = \int_0^1 x f(x) dx = \frac{1}{2} \implies \mu = \frac{100 * \frac{1}{2}}{100} = \frac{1}{2}$

$E(X^2) = \int_0^1 x^2 f(x) dx = \frac{1}{3}$

$Var(X) = E(X^2) - E(X)^2 = \frac{1}{12} \implies \sigma^2 = \frac{100 * \frac{1}{12}}{100^2} = \frac{1}{1200}$

$\bar{X} \in [\mu - 2\sigma, \mu + 2\sigma] \implies \boxed{\bar{X} \in [0.44, 0.56]}$

Problem 5

(a) I liked having recordings of lectures as it helps me review the material before exams. However, not having the settings of classrooms makes it a bit harder to focus during Zoom lectures.

(b) They do a good job of explaining the content. The notes are very helpful as well.

(c) I watch the lectures and the look over the slides/notes occasionally.

(d) Some of the problems feel a bit repetitive but given how much the homeworks are weighted, I would say the workload is reasonable.