Econ 6190 Problem Set 0:
Review of Key Statistical Concepts

Fall 2024

List of key statistical concepts

1. Law of total probability and Bayes rule
2. random variable: its cdf, pdf/pmf, expectation, variance and quantile.
3. random vector: joint cdf, joint pdf/pmf, marginal pdf/pmf
4. conditioning: conditional distribution of a random variable, conditional pdf/pmf, conditional expectation, conditional variance, independence

Practice questions

1. [Hong 2.33] Police plan to enforce speed limits by using radius traps at 4 different locations within the city limits. The radar traps at each of the locations L1, L2, L3, and L4 are operated 40%, 30%, 20%, and 30% of the time, and if a person who is speeding on his way to work has probabilities of 0.2, 0.1, 0.5 and 0.2, respectively, of passing through these locations, what is the probability that he will receive a speeding ticket? Give your reasoning.

2. [Hansen 1.16] Suppose that the unconditional probability of a disease is 0.0025. A screening test for this disease has a detection rate of 0.9, and has a false positive rate of 0.01. Given that the screening test returns positive, what is the conditional probability of having the disease?

3. [Hong 3.5] Suppose $X = X_1$ with probability p and $X = X_2$ with probability $1 - p$, where $p \in (0, 1)$, X_1 and X_2 are random variables with CDF's $F_1(x)$ and $F_2(x)$, respectively. Find the CDF of X.

4. [Hong 3.9] A grocery store sells X hundred kilograms of rice every day, where the distribution of X is of the following form:

$$F(x) = \begin{cases} 0, & x < 0 \\ kx^2, & 0 \leq x < 3, \\ k(-x^2 + 12x - 3), & 3 \leq x < 6, \\ 1 & x \geq 6. \end{cases}$$

Suppose this grocery store's total sales of rice do not reach 600 kilogram on any given day.

(a) Find the value of k.

(b) What is the probability that the store sells between 200 and 400 kilograms of rice next Thursday?

(c) What is the probability that the store sells over 300 kilograms of rice next Thursday?

(d) We are given that the store sold at least 300 kilograms of rice last Friday. What is the probability that it did not sell more than 400 kilograms on that day?

5. [Hong 5.54] Suppose (X, Y) have a joint pdf

$$f_{XY}(x, y) = \begin{cases} xe^{-y}, & \text{if } 0 < x < y < \infty, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the conditional pdf $f_{Y|X}(y|x)$ of Y given $X = x$.

(b) Find the conditional mean $\mathbb{E}(Y|X = x)$.

(c) Find the conditional variance $\text{var}(Y|X = x)$.

(d) Are X and Y independent? Give your reasoning.

6. [Hansen 4.2] Let $f(x, y) = x + y$ for $0 \leq x \leq 1$ and $0 \leq y \leq 1$ (and zero else where).

(a) Verify that $f(x, y)$ is a valid density function.

(b) Find the marginal density of X.

(c) Find $\mathbb{E}[Y]$, $\text{var}[X]$, $\mathbb{E}[XY]$ and $\text{corr}(X, Y)$.

(d) Find the conditional density of Y given $X = x$.

(e) Find $\mathbb{E}[Y|X = x]$.

1. A : the person receives a speeding ticket

A^c : doesn't receive

$$P(A) = 1 - P(A^c)$$

$$= 1 - \sum_{k=1}^{4} P(A^c \mid L_k) \cdot P(L_k)$$

$$= 1 - [(1-0.4)\cdot 0.2 + (1-0.3)\cdot 0.1 + (1-0.2)\cdot 0.5 + (1-0.3)\cdot 0.2]$$

$$= 1 - 0.73$$

$$= 0.27$$

2. event D : a disease

P : screen test positive

$$P(D \mid P) = \frac{P(D \cap P)}{P(P)} = \frac{P(P \mid D)\cdot P(D)}{P(P \mid D)\cdot P(D) + P(P \mid D^c)\cdot P(D^c)}$$

$$= \frac{0.9 \times 0.0025}{0.9 \times 0.0025 + 0.01 \times 0.9975}$$

$$= 0.184$$

3. we know $X = \begin{cases} X_1 & \text{with prob } p \\ X_2 & \text{with prob } 1-p \end{cases}$

$$F_X(x) = P(X \le x)$$

$$= P(X \le x \mid X = X_1)P(X = X_1) + P(X \le x \mid X = X_2)P(X = X_2)$$

$$= P(X_1 \le x)p + P(X_2 \le x)(1-p)$$

$$= pF_1(x) + (1-p)F_2(x).$$

4. (a). Since the grocery store's total sales of rice do't reach 600 kilogram on any given day.

then $P(X = 6) = 0$,

$$\lim_{x \to 6} F(x) = \lim_{x \to 6} k(-x^2 + 12x - 3) = 1$$

then $33k = 1$

we have $k = 1/33$

(b). $P(2 \leq x \leq 4) = P(x \leq 4) - P(X < 2)$

$= P(X \leq 4) - P(X \leq 2)$

$= k(-4^2 + 12 \cdot 4 - 3) - k \cdot 2^2$

$= 25k = 25/33$

(c). $P(X > 3) = 1 - P(X \leq 3)$

$= 1 - 24k = 9/33 = 3/11$

(d). $P(X \geq 3) = 1 - P(X < 3)$

$= 1 - 9k = 24/33$

[Note: $P(X = 3) = 15/33$]

$$P(X \leq 4 \mid X \geq 3) = \frac{P(3 \leq X \leq 4)}{P(X \geq 3)}$$

$$= \frac{P(X \leq 4) - P(X < 3)}{P(X \geq 3)}$$

$$= \frac{29k - 9k}{24k} = 5/6.$$

5. (a). $f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dy$

$= \int_x^{+\infty} x e^{-y}\, dy = x e^{-x}$ if $x > 0$

then $f_{Y|X}(y|x) = \dfrac{f_{XY}(x, y)}{f_X(x)}$

$= \begin{cases} e^{x-y} & \text{if } y > x > 0 \\ 0 & \text{otherwise} \end{cases}$

(b). $E(Y|X) = \int_x^{+\infty} y e^{x-y}\, dy$

$= e^x \cdot \int_x^{+\infty} y e^{-y}\, dy$

$= e^x \cdot \{ y(-e^{-y}) \Big|_x^{+\infty} + \int_x^{+\infty} e^{-y}\, dy$

$= e^x \{ x e^{-x} + e^{-x} \} = x + 1$

(c). $E(Y^2 | X) = \int_x^{+\infty} y^2 e^{x-y}\, dy$

$= e^x \int_x^{+\infty} y^2 e^{-y}\, dy$

$= e^x \{ y^2(-e^{-y}) \Big|_x^{+\infty} + \int_x^{+\infty} 2y e^{-y}\, dy \}$

$= e^x \{ x^2 e^{-x} + 2e^{-x}(x+1) \}$

$= x^2 + 2x + 2$

$Var(Y|X) = E(Y^2|X) - \left(E(Y|X)\right)^2 = x^2 + 2x + 2 - (x+1)^2$

$= 1$

(d). No. $E(Y|X) = X + 1 \neq EY$

$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dx$

$= \int_0^y x e^{-y}\, dx = \begin{cases} \frac{1}{2} y^2 e^{-y} & y > 0 \\ 0 & \text{otherwise} \end{cases}$

EY sholild be independent of X.

6. (a).

$$\int_0^1 \int_0^1 (x+y)\,dx\,dy$$

$$= \int_0^1 (\tfrac{1}{2} + y)\,dy = \tfrac{1}{2} + \tfrac{1}{2} = 1$$

moreover, since when $x \in [0,1]$, $y \in [0,1]$

$x+y \geq 0$, then it is valid

(b). marginal density of X

$$\int_0^1 (x+y)\,dy = \tfrac{1}{2} + x$$

$$f_X(x) = \begin{cases} \tfrac{1}{2} + x & \text{if } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

(c). $f_Y(y) = \tfrac{1}{2} + y$

$$EY = \int_0^1 y \cdot (\tfrac{1}{2} + y)\,dy$$

$$= \int_0^1 (\tfrac{1}{2}y + y^2)\,dy = (\tfrac{1}{4}y^2 + \tfrac{1}{3}y^3)\Big|_0^1$$

$$= \tfrac{1}{4} + \tfrac{1}{3} = \tfrac{7}{12}$$

$$\text{Var}\,X = EX^2 - (EX)^2$$

$$= \int_0^1 x^2(\tfrac{1}{2} + x)\,dx - (\tfrac{7}{12})^2$$

$$= (\tfrac{1}{6}x^3 + \tfrac{1}{4}x^4)\Big|_0^1 - (\tfrac{7}{12})^2$$

$$= \tfrac{1}{6} + \tfrac{1}{4} - \tfrac{49}{144} = \frac{24+36-49}{144} = \frac{11}{144}$$

$$EXY = \int_0^1 \int_0^1 xy(x+y)\,dx\,dy$$

$$= \int_0^1 y\left[\int_0^1 x^2 + xy\,dx\right]dy$$

$$= \int_0^1 y \cdot (\tfrac{1}{3} + \tfrac{1}{2}y)\,dy = \tfrac{1}{3}$$

$$\operatorname{corr}(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{Var} X \cdot \operatorname{Var} Y}} = \frac{EXY - EX \cdot EY}{\operatorname{Var} X}$$

(X & Y are symmetric)

$$= \frac{\frac{1}{3} - \left(\frac{7}{12}\right)^2}{11/144} = -1/11$$

(d). $$f_{Y|X}(y \mid X = x) = \frac{f(x, y)}{f_X(x)} = \frac{x+y}{x+\frac{1}{2}} \qquad y \in [0, 1]$$

(e). $$E(Y \mid X = x) = \int_0^1 y \cdot \frac{x+y}{x+\frac{1}{2}} \, dy$$

$$= \frac{1}{x+\frac{1}{2}} \int_0^1 [xy + y^2] \, dy$$

$$= \frac{1}{x+\frac{1}{2}} \cdot \left(\frac{1}{2}x + \frac{1}{3}\right)$$

$$= \frac{1}{2} + \frac{1}{12x+6}$$

[check: $$EY = E_X E(Y \mid X = x)$$

$$= E_X\left[\frac{1}{2} + \frac{1}{12x+6}\right]$$

$$= \int_0^1 \left(\frac{1}{2} + \frac{1}{12x+6}\right) \cdot \left(x + \frac{1}{2}\right) dx$$

$$= \frac{1}{12} + \frac{1}{4} + \frac{1}{4} = \frac{7}{12}$$ #]