Lecture 10

(5 September 2024)

Recap

RVs $X_1, X_2, \dots, X_n$ are independent if

$$p_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$$

$$\forall\, x_1, x_2, \dots, x_n.$$

Linearity of Expectation:

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

If $X_1, X_2, \dots, X_n$ are independent RVs

$$\operatorname{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{var}(X_i).$$

$$p_{X|A}(x) = P(X=x \mid A) = P(X=x \cap A)/P(A)$$

$$p_{X|Y}(x|y) = p_{XY}(x,y)/p_Y(y)$$

$$p_{XY}(x,y) = p_{X|Y}(x|y)\, p_Y(y) = p_{Y|X}(y|x)\, p_X(x)$$

Conditional Expectation

The conditional expectation of X given $Y=y$ is defined as

$$E[X|Y=y] = \sum_x x\, p_{X|Y}(x|y).$$

Similarly for an event A with $P(A)>0$

$$E[X|A] = \sum_x x\, p_{X|A}(x).$$

Theorem. $E[g(X)|A] = \sum_x g(x)\, p_{X|A}(x).$

The proof is exactly similar to the proof of $E[g(X)] = \sum_x g(x)\, p_X(x).$

Total Expectation Theorem:

If the events $A_1, A_2, \ldots, A_n$ form a partition of the sample space with $P(A_i)>0$ for all i then

$$E[X] = \sum_{i=1}^{n} P(A_i)\, E[X|A_i].$$

Proof:

$$\sum_{i=1}^{n} P(A_i) E[X \mid A_i]$$

$$= \sum_{i=1}^{n} P(A_i) \sum_{x} x \, p_{X \mid A_i}(x)$$

$$= \sum_{x} \sum_{i=1}^{n} x \, \frac{\cancel{P(A_i)} \, P(X = x \cap A_i)}{\cancel{P(A_i)}}$$

$$= \sum_{x} x \, P(\{X = x\} \cap \Omega)$$

$$= \sum_{x} x \, p_X(x) = E[X].$$

Taking $\{Y = y\}_{y \in Y}$ as the partition of Ω the total expectation theorem gives

$$E[X] = \sum_{y} p_Y(y) E[X \mid Y = y].$$

Conditional Expectation as a RV:

Let $\phi(y) = E[X \mid Y=y]$

Y is a RV, $\phi(Y)$ is also a RV.

We denote $\phi(Y) \equiv E[X \mid Y]$.

$E[X \mid Y]$ is a function of RV Y.

$$E[\phi(Y)] = E[E[X \mid Y]]$$

$$= \sum_y p_Y(y)\,\phi(y)$$

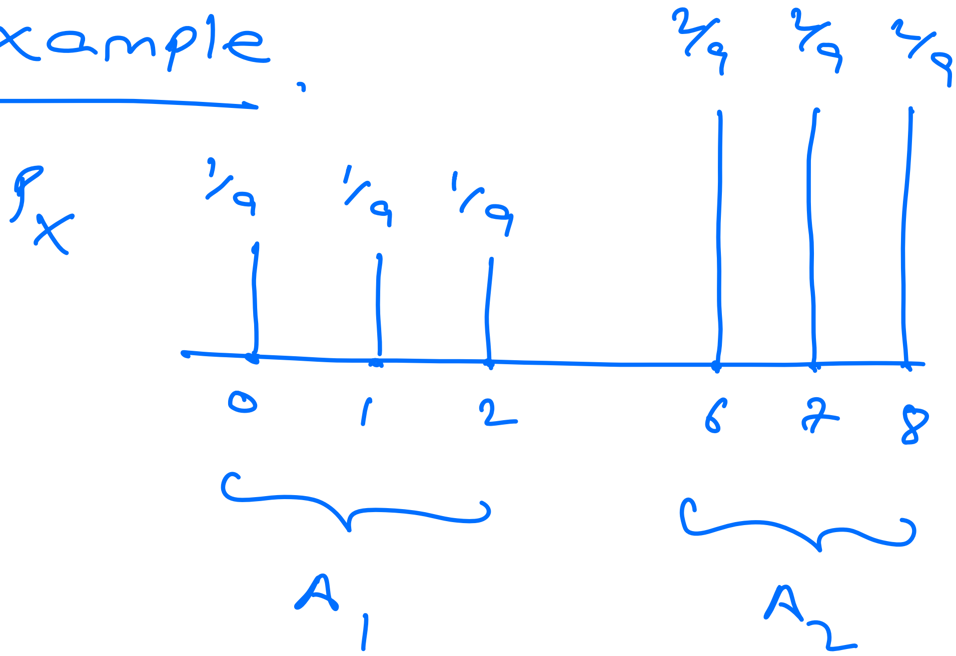
$$= \sum_y p_Y(y)\,E[X \mid Y=y]$$

$$= E[X]$$

(by total expectation theorem).

Thus $E[E[X \mid Y]] = E[X]$ — called the law of iterated expectations.

Example.



$$P(A_1) = \frac{1}{3} \qquad P(A_2) = \frac{2}{3}$$

$$E[X \mid A_1] = 0 \cdot \tfrac{1}{3} + 1 \cdot \tfrac{1}{3} + 2 \cdot \tfrac{1}{3}$$

$$= 1$$

$$E[X \mid A_2] = \frac{6+7+8}{3} = 7$$

$$E[X] = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 7 = 5.$$

Conditional Independence

Two random variables X and Y are conditionally independent given an event A if

$$P_{XY|A}(x,y) = P_{X|A}(x)\, P_{Y|A}(y) \quad \forall x,y.$$

Note that $P_{XY|A}(x,y) = \dfrac{P(X=x \cap Y=y \cap A)}{P(A)}$.

Exercise, Consider the following P_{XY}.

$x \backslash y$	1	2	3	4
1	0	1/20	0	0
2	0	1/20	3/20	1/20
3	2/20	4/20	1/20	2/20
4	1/20	2/20	2/20	0

Are X and Y independent?

Are X and Y conditionally independent given $A = \{X \geq 3,\ Y \leq 2\}$?

Conditional Variance

Recall variance of X,

$$Var(X) = E[(X - E[X])^2].$$

$\sigma_X = \sqrt{Var(X)}$ is called standard deviation.

$$Var(X \mid Y=y) = \psi(y)$$

$$= E\left[\left(X - E[X \mid Y=y]\right)^2 \mid Y=y\right]$$

$$= E\left[X^2 + E[X \mid Y=y]^2 - 2X E[X \mid Y=y] \mid Y=y\right]$$

$$= E[X^2 \mid Y=y] + E[X \mid Y=y]^2$$
$$- 2E[X \mid Y=y]^2$$

$$= E[X^2 \mid Y=y] - E[X \mid X=y]^2.$$

$\mathrm{Var}(X|Y) = \varphi(Y)$ is a RV.

Theorem (Law of total variance).

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

Proof.

$$\mathrm{Var}(X) = E[X^2] - E[X]^2$$

$$= E[E[X^2|Y]] - E[X]^2$$

$$= E[\mathrm{Var}(X|Y) + E[X|Y]^2] - E[E[X|Y]]^2$$

$$= E[\mathrm{Var}(X|Y)] + E[E[X|Y]^2] - E[E[X|Y]]^2$$

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

Problems

1) Consider Bin$(n=100, p=0.01)$. Find the probability of 5 successes in 100 trials, approximately [Note n is large].

Sol:- We use Poisson's approximation.

$$\lambda = np = 1.$$

$$P_X(5) = \frac{e^{-1} 1^5}{5!} = \frac{e^{-1}}{5!} \approx 0.003.$$

2) Let X be a geometric random variable with parameter p.

$$P(X>n) = \sum_{i=n+1}^{\infty} (1-p)^{i-1} p$$

$$= \cancel{p} \cdot \frac{(1-p)^n}{\cancel{p}} = (1-p)^n$$

$$P(X>m+n \mid X>m) \quad [m, n \in \mathbb{N}]$$

$$= \frac{P(X>m+n,\ X>m)}{P(X>m)} = \frac{P(X>m+n)}{P(X>m)}$$

$$= \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = P(X>n).$$

$$\therefore P(X>m+n \mid X>m) = P(X>n).$$

This is called memoryless property of the geometric random variable.

Q) You are allowed to take a certain test three times and your final score will be the maximum of the test scores.

$$X = \max\{x_1, x_2, x_3\},$$ where

x_1, x_2, x_3 are the three test scores and X is the final score. Assume that your score in each test takes one of the values from 1 to 10 with equal probability $1/10$ - independently of the scores in other tests. Find the PMF of the final score.

Sol:- $F_X(k) = P(X \leq k)$

$= P(\max\{x_1, x_2, x_3\} \leq k)$

$= P(x_1 \leq k, x_2 \leq k, x_3 \leq k)$

$= P(X_1 \leq k)\, P(X_2 \leq k)\, P(X_3 \leq k)$

[as X_1 X_2 and X_3 are independent]

$= F_{X_1}(k)\, F_{X_2}(k)\, F_{X_3}(k)$

$P_X(k) = F_X(k) - F_X(k-1)$

$= F_{X_1}(k)\, F_{X_2}(k)\, F_{X_3}(k) - F_{X_1}(k-1)\, F_{X_2}(k-1)\, F_{X_3}(k-1)$

$= \dfrac{k^3 - (k-1)^3}{1000}$, $k = 1, 2, \ldots$