CSE 107 10-22-24

- midterm 1 : Thur. 10/24
  - test : 9:50 - 10:55
  - break : 10:55 - 11:00
  - lecture : 11:00 - 11:25

2.7 Independence of r.v.s

Defn

We say r.v.s X, Y are independent iff events

$$\{X = x\} \text{ and } \{Y = y\}$$

are independent for all $x \in \text{range}(X)$ and $y \in \text{range}(Y)$, i.e.

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$$

i.e.

$$\boxed{p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)}$$

for all possible x, y.

If $P_Y(y) > 0$ for all $y \in \text{range}(Y)$,

then

$$\frac{P_{X,Y}(x,y)}{P_Y(y)} = P_X(x)$$

i.e.

$$\boxed{P_{X|Y}(x|y) = P_X(x)}$$

for all poss. x, y

As usual, extend defn. to conditional Probability law: $P(\cdot \mid A)$ where $P(A) > 0$

<u>Defn</u>

We say X, Y are <u>conditionally independent given A</u> iff

$$P(X=x, Y=y \mid A) = P(X=x \mid A) \cdot P(Y=y \mid A)$$

for all $x \in \text{range}(X)$, $y \in \text{range}(Y)$.

i.e.

$$\boxed{P_{X,Y|A}(x,y) = P_{X|A}(x) \cdot P_{Y|A}(y)}$$

Theorem

If X, Y are independent, then

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

Proof

$$E[X \cdot Y] = \sum_x \sum_y xy \, p_{X,Y}(x,y)$$

$$= \sum_x \sum_y x \cdot y \cdot p_X(x) \cdot p_Y(y)$$

$$= \left(\sum_x x p_X(x) \right) \cdot \left(\sum_y y p_Y(y) \right)$$

$$= E[X] \cdot E[Y].$$

∎

Theorem

If X, Y are independent r.v.s, then so are $g(X)$ and $h(Y)$

See Problem #44 on p. 134 of text.

Thus

$$E[g(X) \cdot h(Y)] = E[g(X)] \cdot E[h(Y)].$$

Theorem

If X, Y are independent, then

$$Var(X+Y) = Var(X) + Var(Y).$$

Proof

let X, Y be independent, and let

$$U = X - E[X], \quad V = Y - E[Y].$$

Then

$$E[U] = E[X - E[X]] = E[X] - E[X] = 0$$

and

$$E[V] = \cdots = 0.$$

Also

$$Var(U) = Var(X)$$
$$Var(V) = Var(Y).$$

$\Rightarrow 0$

$$U+V=(X-E[X])+(Y-E[Y])$$

$$=(X+Y)-E[X+Y]$$

Thus

$$Var(X+Y)=E[(U+V)^2]$$

$$=E[U^2+2UV+V^2]$$

$$=E[U^2]+2E[UV]+E[V^2]$$

$$=E[U^2]+2\cancel{E[U]}^{0}\cdot\cancel{E[V]}^{0}+E[V^2]$$

$$=Var(X)+Var(Y).$$

$\blacksquare$

we can do this for 3, 4, 5... r.v.s. For instance, let X, Y, Z be r.v.s. They are independent iff

$$p_{X,Y,Z}(x,y,z) = p_X(x) \cdot p_Y(y) \cdot p_Z(z).$$

If X, Y, Z are ind., then so are $f(X), g(Y), h(Z)$ (or any other functional combinations of X, Y, Z)

we have $E[XYZ] = E[X]E[Y]E[Z]$

and $Var(X+Y+Z) = Var(X) + Var(Y) + Var(Z)$.

Variance of Binomial r.v.

Recall, if X is Binomal with Param. n, p, then

$$X = \underbrace{X_1 + X_2 + \cdots + X_n}_{\text{independent, identically distributed (i.i.d) Bernoulli r.v.s}}$$

Thus

$$\begin{aligned} Var(X) &= Var(X_1) + \cdots + Var(X_n) \\ &= p(1-p) + \cdots + p(1-p) \\ &= np(1-p). \end{aligned}$$

Ex.

suppose we flip a weighted coin, but we don't know

$$P(\text{head}) = p$$

How can we estimate p?

we compute the sample mean

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

where $X_i = \begin{cases} 1 & \text{if } i^{th} \text{ flip is heads} \\ 0 & \text{if } i^{th} \text{ flip is tails} \end{cases}$

we have $E[X_i] = p$, we have

$$E[S_n] = E\left[\tfrac{1}{n}X_1 + \tfrac{1}{n}X_2 + \cdots + \tfrac{1}{n}X_n\right]$$

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

$$= \sum_{i=1}^{n} \tfrac{1}{n}\cdot p = n\cdot\tfrac{1}{n}\cdot p = p$$

so S_n is an estimate for p.

since X_i are ind., so are the $\frac{1}{n}\cdot X_i$. Hence

$$\mathrm{Var}(S_n) = \mathrm{Var}\left(\sum_{i=1}^{n}\tfrac{1}{n}X_i\right)$$

$$= \sum_{i=1}^{n}\mathrm{Var}\left(\tfrac{1}{n}X_i\right)$$

$$= \sum_{i=1}^{n} \frac{1}{n^2} \mathrm{Var}(X_i)$$

$$= n \cdot \frac{1}{n^2} p(1-p)$$

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

$\therefore \mathrm{Var}(S_n) \longrightarrow 0$ as $n \to \infty$

since $E[S_n] = p$, we have

$$S_n \longrightarrow p \quad \text{as } n \to \infty.$$

<u>Note</u>: Even if $X_i \overset{\text{"}}{X}$ are not Bernoulli,

the same calculation gives

$$Var(S_n) = \frac{Var(X)}{n}$$

Provided X_i are i.i.d. Since

$$E[S_n] = E[X_i],$$

so S_n is a good approx. to mean $E[X]$.

Ex. (#38 on P. 132, ch 2)

Alice passes through 4 traffic lights, each equally likely to be Red or Green independently.

(a) find PMF, mean & variance of # of Red lights encountered.

Solution

Let $X = \#$ red lights. Then X is binomial with param.

$n = 4 \qquad p = \frac{1}{2}$

so

$$P_X(k) = \binom{4}{k}\left(\tfrac{1}{2}\right)^k\left(\tfrac{1}{2}\right)^{4-k} \qquad k=0,1,2,3,4$$

$$= \binom{4}{k}\left(\tfrac{1}{2}\right)^4$$

$$= \frac{4!}{k!(4-k)!}\cdot\frac{1}{16}$$

$$= \frac{24}{16\cdot k!(4-k)!}$$

$$= \boxed{\frac{3}{2k!(4-k)!} \qquad k=0,1,2,3,4}$$

$$E[X] = np = 4\cdot\tfrac{1}{2} = \boxed{2}$$

$$\mathrm{Var}(X) = n\cdot p(1-p) = 4\cdot\tfrac{1}{2}\cdot\tfrac{1}{2} = \boxed{1}$$

(b) suppose Alice is delayed by 2 minutes for each red light. Find the Variance of her Commute time.

Solution

Let T be her commute time. Then

$$T = T_0 + 2X$$

where T_0 is her commute time with no red lights.

so

$$\mathrm{Var}(T) = \mathrm{Var}(T_0 + 2X)$$

$$= \mathrm{Var}(2X)$$

$$= 2^2 \cdot \mathrm{Var}(X)$$

$$= 4 \cdot 1 = \boxed{4}$$