2.7 Independence

Defn

Two r.v.s X, Y are called **independent** iff for all x, y the events

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

are independent.

i.e.

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

i.e.

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

if $P_Y(y) > 0$ for all y, this is eq. to

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

i.e.

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

we can extend all this to the conditional setting: for $P(A)>0$

$$P(\cdot) \rightsquigarrow P(\cdot \mid A)$$

Defn

we say X, Y are conditionally independent given A iff

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

i.e.

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

<u>note</u>: if $P_{Y|A}(y) > 0$ for all y

equiv. to

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

Recall

$$\boxed{\text{independence}} \;\begin{matrix} \not\Rightarrow \\ \not\Leftarrow \end{matrix}\; \boxed{\begin{matrix}\text{conditional} \\ \text{independence}\end{matrix}}$$

see fig. 2.15 on p. 111 of text.

<u>Note</u>: if X, Y are independent, then

$$\boxed{E[X \cdot Y] = E[X] \cdot E[Y]}$$

<u>Proof</u>.

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

$$= \sum_x \sum_y xy\, P_X(x) \cdot P_Y(y)$$

$$= \left(\sum_x x P_X(x)\right) \cdot \left(\sum_y y P_Y(y)\right)$$

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

■

Note:

if X, Y are independent, then so are $g(X), h(Y)$.

see Prob. #44 on P. 134 of text.

Consequence:

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

Also: if X, Y are independent, then

$$\boxed{Var(X+Y) = Var(X) + Var(Y)}$$

Proof.

Let X, Y be ind. r.v.s, set

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

so

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

and

$$E[V] = \qquad\qquad = 0$$

Also

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

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

Hence

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

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

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

$$= E[U^2] + 2\underbrace{\cancel{E[U]}\cdot\cancel{E[V]}}_{\text{by ind.}} + E[V^2]$$

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

$$= \mathrm{Var}(X) + \mathrm{Var}(Y)$$

$\blacksquare$

all this can be done for 3 or more r.v.s :

we call X, Y, Z independent iff

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

for all x, y, z. In this case any fcns of X, Y, Z are independent.

Ex $g(X,Y)$ and $h(Z)$

Similarly: if X, Y, Z are ind.

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

likewise for any number of r.v.s

Variance of Binomial

Recall: if X is Bernoulli, then

$$E[X] = p$$

$$Var(X) = p(1-p)$$

Recall: Binomial Y is a sum of ind. Bernoulli

$$Y = X_1 + X_2 + \cdots + X_n$$

all with same parameter p. So

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)$$

$$= p(1-p) + p(1-p) + \cdots + p(1-p)$$

$$= np(1-p).$$

Variance of Poisson

let X be Poisson w/ param $\lambda > 0$

recall $E[X] = \lambda$. will show

$Var(X) = \lambda$.

Proof.

$$E[X^2] = \sum_{k=\cancel{0}1}^{\infty} k^2 \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$$= \lambda \cdot \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^{k-1}}{(k-1)!}$$

$$= \lambda \cdot \sum_{k=0}^{\infty} (k+1) e^{-\lambda} \cdot \frac{\lambda^k}{k!} \qquad k \leftarrow k+1$$

$$= \lambda \left(\sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \right)$$

$$= \lambda \left(E[X] + 1 \right)$$

$$= \lambda(\lambda+1) = \lambda^2 + \lambda$$

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

$$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2}$$

$$= \lambda .$$

<u>Ex.</u>

suppose we have a weighted coin but we don't know $P(\text{head}) = p$.

How can we estimate p?

flip coin n times. let

$$X_i = \begin{cases} 1 & \text{if head} \\ 0 & \text{if tail} \end{cases}$$

Define <u>Sample mean</u> S_n

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

(also called <u>relative frequency</u>)

i.e.

$$S_n = \sum_{i=1}^{n} \frac{1}{n} X_i$$

Know $E[X_i] = p$. Then

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

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

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

so S_n approximates p.

but how good?

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

hence

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

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

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

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

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

$\therefore \ Var(S_n) \to 0$ as $n \to \infty$

so S_n gets better as $n \to \infty$.

<u>note</u>: even if X_i are not Bernoulli, as long as they're ind. we have

similar calculation

$$Var(S_n) = \ldots\ldots$$
$$= \ldots\ldots$$
$$= \frac{Var(X)}{n} \to 0$$

as $n \to \infty$.

Ex. (#38 on p.132)

Alice passes through 4 traffic lights on way to work, each eq. likely to be red or green.

(a) find PMF, mean and variance of # of red lights X.