Independence of Random Variables.

Def. R.v's X and Y are independent if their joint distribution function factors into the product of their marginal distribution functions:

$$F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$$

Theorem. Suppose X and Y are jointly continuous r.v's. X and Y are independent iff given two densities for X and Y their product is the joint density for the pair (X,Y), i.e.

$$f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$$

Proof:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_X(x)\,F_Y(y)$$
$$= \frac{\partial}{\partial x} F_X(x) \frac{\partial}{\partial y} F_Y(y) = f_X(x)\,f_Y(y) \quad \blacksquare$$

Note: If X and Y are independent, and $Z = g(X)$, $W = h(Y)$ then Z and W are also independent.

Ex. Let X, Y be discrete r.v's

$$P_{X,Y}(x,y) = \frac{1}{x!\,y!}\lambda^x \mu^y e^{-(\lambda+\mu)} \qquad x, y = 0, 1, 2, \ldots$$

Are X and Y independent? What are their marginal distributions?

$$P_X(x) = \sum_y \frac{1}{x!\,y!}\lambda^x \mu^y e^{-(\lambda+\mu)}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} \underbrace{\sum_y \frac{\mu^y e^{-\mu}}{y!}}_{=1} = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$P_Y(y) = \frac{\mu^y e^{-\mu}}{y!}$$

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

Ex.

$$f_{X,Y}(x,y) = \begin{cases} 4(x+y^2), & x,y > 0,\ x+y \le 1 \\ 0, & \text{ow} \end{cases}$$

Are X and Y independent?

$$f_X(x) = \int_0^{1-x} 4(x+y^2)\,dy = \left(4xy + \frac{4y^3}{3}\right)\Bigg|_0^{1-x}$$

$$= \cancel{4x} - \cancel{4x^2} + \frac{4}{3}(1 - \cancel{3x} + \cancel{3x^2} - x^3)$$

$$= \frac{4}{3}(1 - x^3)$$

$$\boxed{f_X \cdot f_Y \neq f_{X,Y}}$$

$$f_Y(y) = \int_0^{1-y} 4(x+y^2)\,dx = \left(4\frac{x^2}{2} + 4y^2 x\right)\Bigg|_0^{1-y}$$

$$= 2(1-y)^2 + 4y^2(1-y) = 2 - 4y + 6y^2 - 4y^3$$

Conditional Densities.

Def. If X, Y are jointly distributed continuous r.v's, the <u>conditional density function</u> of $Y|X$ is

$$f_{Y|X}(y|x) = \begin{cases} \dfrac{f_{X,Y}(x,y)}{f_X(x)}, & f_X(x) > 0 \\ 0, & \text{ow} \end{cases}$$

If X, Y are independent,

$$f_{Y|X}(y|x) = f_Y(y)$$

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

$$\Rightarrow \quad f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x)\, f_X(x)\, dx$$

Ex.

$$f_{X,Y}(x,y) = \begin{cases} \lambda^2 e^{-\lambda y}, & 0 \le x \le y < \infty \\ 0, & \text{ow} \end{cases}$$

Find $f_{X|Y}$ and $f_{Y|X}$.

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

$$f_Y(y) = \int_0^{y} \lambda^2 e^{-\lambda y}\, dx = \lambda^2 x e^{-\lambda y}\Big|_0^{y} = \lambda^2 y e^{-\lambda y}$$

$$f_{X|Y} = \frac{\lambda^2 e^{-\lambda y}}{\lambda^2 y e^{-\lambda y}} = \frac{1}{y}$$

$$f_{Y|X} = \frac{\lambda^2 e^{-\lambda y}}{\lambda e^{-\lambda x}} = \lambda e^{-\lambda (y-x)}$$

Properties of Expectations;

- $E(aX+bY) = aE(X) + bE(Y)$, $a, b \in \mathbb{R}$

Pf: $E(aX+bY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (ax+by)\, f_{X,Y}(x,y)\, dx\, dy$

$= a \iint x\, f_{X,Y}(x,y)\, dx\, dy + b \iint y\, f_{X,Y}(x,y)\, dx\, dy$

$= a \int x\, f_X(x)\, dx + b \int y\, f_Y(y)\, dy = aE(X) + bE(Y)$

- Given X, Y are independent then

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

$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{X,Y}(x,y)\, dx\, dy = \iint xy\, f_X(x) f_Y(y)\, dx\, dy$

$= \int x\, f_X(x)\, dx \int y\, f_Y(y)\, dy = E(X)E(Y)$ ∎

Covariance.

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

$= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2$

$= Var(X) + Var(Y) + 2\, cov(X,Y)$

Def. For r.v's X, Y with $E(X), E(Y) < \infty$, the **covariance** of X and Y is

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

Claim: $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$

Pf. $E[(X-E(X))(Y-E(Y))]$

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

$$= E(XY) - E(X)E(Y) - \cancel{E(X)E(Y)} + \cancel{E(X)E(Y)}$$

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

Note: X, Y are indep., then $\mathrm{Cov}(X,Y) = 0$

Ex.

$y \backslash x$	-1	0	1	$P_Y(y)$
-1	$1/8$	$1/8$	$1/8$	$3/8$
0	$1/8$	0	$1/8$	$2/8$
1	$1/8$	$1/8$	$1/8$	$3/8$
$P_X(x)$	$3/8$	$2/8$	$3/8$	

Find $\mathrm{Cov}(X,Y)$. Are X, Y independent?

$$E(X) = -1 \cdot \frac{3}{8} + 0 \cdot \frac{2}{8} + 1 \cdot \frac{3}{8} = 0 = E(Y)$$

$$E(XY) = \sum\sum xy\, p_{XY}(x,y) = (-1)(-1)\frac{1}{8} + (-1)1\frac{1}{8} + 1\cdot(-1)\frac{1}{8} + 1\cdot 1\frac{1}{8}$$

$$= 0$$

$$\mathrm{Cov}(X,Y) = 0$$

$$p_{XY}(-1,-1) = \frac{1}{8} \neq p_X(-1)\,p_Y(-1) = \frac{3}{8}\cdot\frac{3}{8} = \frac{9}{64}$$

$\Rightarrow X, Y$ are NOT independent

Important Facts:

- Independence of X, Y implies $Cov(X,Y) = 0$, but NOT vice versa.

- If X, Y are independent, then $Var(X \pm Y) = Var(X) + Var(Y)$

- If X, Y are not independent, then

$$Var(X \pm Y) = Var(X) + Var(Y) \pm 2\, Cov(X,Y)$$

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

<u>Ex</u>. $Y \sim Bin(n, p)$. $Var(Y) = ?$

$X_i \sim iid$ Bernoulli (p)

iid = independent identically distributed

$E(X_i) = p$, $Var(X_i) = pq$

$$Y = \sum_{i=1}^{n} X_i, \quad Var(Y) = \sum_{i=1}^{n} Var(X_i)$$

$$= npq$$

Properties of Covariance:

For r.v's X, Y, Z and constants $a, b, c, d \in \mathbb{R}$

(1) $Cov(aX+b, cY+d) = ac\, Cov(X,Y)$

$= E[acXY + adX + bcY + bd] - (aE(X)+b)(cE(Y)+d)$

$= acE(XY) - acE(X)E(Y) = ac\, Cov(X,Y)$

(2) $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$

$= E[XZ + YZ] - E(X+Y)E(Z) =$

$= E(XZ) + E(YZ) - E(X)E(Z) - E(Y)E(Z)$

$= Cov(X,Z) + Cov(Y,Z)$

(3) $Cov(X,Y) = cov(Y,X)$

Correlation:

Def. For r.v's X, Y the correlation of X and Y is

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}, \quad V(X), V(Y) \neq 0$$

Claim: $\rho(aX+b, cY+d) = \rho(X,Y)$

$$= \frac{Cov(aX+b, cY+d)}{\sqrt{Var(aX+b)\,Var(cY+d)}} = \frac{\cancel{ac}\, Cov(X,Y)}{\cancel{ac}\sqrt{Var(X)\,Var(Y)}}$$

$$= \rho(X,Y)$$

$\Rightarrow$ correlation is scale invariant

Theorem. $-1 \le \rho(X,Y) \le 1$

Proof: $0 \le Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) =$

$= Var\left(\frac{X}{\sigma_X}\right) + Var\left(\frac{Y}{\sigma_Y}\right) + 2\,cov\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right)$

$= \frac{1}{\sigma_X^2} Var(X) + \frac{1}{\sigma_Y^2} Var(Y) + \frac{2\,cov(X,Y)}{\sigma_X \sigma_Y}$

$= 1 + 1 + 2\rho(X,Y) = 2(1+\rho(X,Y))$

$\Rightarrow \rho(X,Y) \ge -1$

$0 \le Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2(1-\rho(X,Y))$

$\Rightarrow \rho(X,Y) \le 1$ $\blacksquare$

Interpretation of Correlation ρ:

If $\rho = 1 \Rightarrow Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0$

$\Rightarrow P\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = c\right) = 1$, c is a const.

$\Leftrightarrow P(Y = a + bX) = 1$, $a = -c\sigma_Y$, $\underbrace{b = \frac{\sigma_Y}{\sigma_X}}_{>0}$

If $\rho = -1 \Rightarrow Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0$

$\Rightarrow P(Y = a + bX) = 1$, $a = c\sigma_Y$, $\underbrace{b = -\frac{\sigma_Y}{\sigma_X}}_{<0}$

So,

Y is an increasing/(decreasing) linear function of X iff $\rho(X,Y) = 1$ (-1)

If X, Y are indep., then $\rho(X,Y) = 0$.

Ex.

$$f_{X,Y}(x,y) = \begin{cases} 3x, & 0 \le y \le x \le 1 \\ 0, & \text{ow} \end{cases}$$

Find Var(X−Y) and $\rho(X,Y)$. $= \dfrac{Cov(X,Y)}{\sqrt{Var X \, Var(Y)}} = \dfrac{3/160}{\sqrt{\frac{3}{80} \cdot \frac{19}{320}}} = \sqrt{3/19}$

Sol'n: $Var(X-Y) = Var(X) + Var(Y) - 2\,Cov(X,Y)$

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

$$f_X(x) = \int_0^x 3x\,dy = 3x^2, \quad 0 \le x \le 1$$

$$f_Y(y) = \int_y^1 3x\,dx = \frac{3}{2}(1-y^2), \quad 0 \le y \le 1$$

$$E(X) = \int_0^1 x \cdot 3x^2\,dx = \frac{3}{4}$$

$$E(Y) = \frac{3}{2}\int_0^1 y(1-y^2)\,dy = \frac{3}{2}\left(\frac{y^2}{2} - \frac{y^4}{4}\right)\Big|_0^1 = \frac{3}{8}$$

$$E(XY) = \int_0^1\int_0^x xy \cdot 3x\,dy\,dx = \frac{3}{2}\int_0^1 x^4\,dx = \frac{3}{10}$$

$$Cov(X,Y) = \frac{3}{10} - \frac{3}{4}\cdot\frac{3}{8} = \frac{3}{10} - \frac{9}{32} = \frac{3}{160}$$

$$Var(X) = E(X^2) - E(X)^2 = \int_0^1 3x^4\,dx - \frac{9}{16}$$

$$= \frac{3}{5} - \frac{9}{16} = \frac{48-45}{80} = \frac{3}{80}$$

$$Var(Y) = \frac{3}{2}\int_0^1 y^2(1-y^2)\,dy - \left(\frac{3}{8}\right)^2 = \frac{19}{320}$$

$$Var(X-Y) = \cancel{\frac{3}{80}} + \frac{19}{320} - 2\cdot\cancel{\frac{3}{160}} = \frac{19}{320}$$

Conditional Expectation.

If X, Y are discrete r.v's, then

$$E[Y \mid X=x] = \sum_{y} y \, P_{Y|X}(y \mid x)$$

$$Var[Y \mid X=x] = \sum_{y} (y - E(Y \mid X=x))^2 P_{Y|X}(y \mid x)$$

$$= E[Y^2 \mid X=x] - [E(Y \mid X=x)]^2$$

In general,

$$E[h(Y) \mid X=x] = \sum_{y} h(y) \, P_{Y|X}(y \mid x)$$

If X, Y are continuous r.v's, then

$$E(Y \mid X=x) = \int_{-\infty}^{\infty} y \, f_{Y|X}(y \mid x) \, dy$$

$$Var(Y \mid X=x) = \int_{-\infty}^{\infty} [y - E[Y \mid X=x]]^2 f_{Y|X}(y \mid x) \, dy$$

$$= E(Y^2 \mid X=x) - (E(Y \mid X=x))^2$$

In general,

$$E(h(Y) \mid X=x) = \int_{-\infty}^{\infty} h(y) f_{Y|X}(y \mid x) \, dy$$

Ex.

$$f_{X,Y}(x,y) = \begin{cases} e^{-y}, & y>0,\ 0\le x\le 1 \\ 0, & ow \end{cases}$$

Find $E(X \mid Y=2)$.

$$f_{X|Y} = \frac{f_{X,Y}}{f_Y} = 1, \quad 0 \le X \le 1$$

$$f_Y(y) = \int_0^1 e^{-y}\,dx = e^{-y}, \quad y>0$$

$$E(X|Y) = \int_0^1 x\, f_{X|Y}(x|y)\,dx$$

$$= \int_0^1 x \cdot 1\,dx = \frac{1}{2}$$

X, Y are indep.

Ex

$$f(x,y) = \begin{cases} \frac{1}{2}, & 0 \le X \le y \le 2 \\ 0, & \text{elsewhere} \end{cases}$$

Find $E(X \mid Y = 1.5)$

$$f_Y(y) = \int_0^y \frac{1}{2}\,dX = \frac{y}{2}, \quad 0 \le y \le 2$$

$$f_{X|Y}(x|y) = \frac{1/2}{y/2} = \frac{1}{y}, \quad 0 \le X \le y$$

$$E(X|Y) = \int_0^y X \cdot \frac{1}{y}\,dX = \frac{y^2}{2}\cdot\frac{1}{y} = \frac{y}{2}$$

$$E(X|Y=1.5) = \frac{1.5}{2} = 0.75$$

Assume $E(Y|X=x)$ exists for every x in the range of X. Then $E(Y|X)$ is a r.v.

<u>Theorem</u>: $E[E(Y|X)] = E(Y) \longrightarrow$ "Law of Total Expectation"

<u>Proof</u>:

$$E[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f_{X,Y}(x,y)\,dx\,dy =$$

$$= \iint y\, f_{Y|X}(y|x)\, f_X(x)\,dy\,dx$$

$$= \int_{-\infty}^{\infty} \underbrace{\left[\int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\,dy\right]} f_X(x)\,dx$$

$$= \int_{-\infty}^{\infty} E(Y|X)\, f_X(x)\,dx = E[E(Y|X)]$$

<u>Ex</u>. Roll a die. Whatever number comes up we toss a coin that many times. What is the expected # of heads?

$Y = \#$ on a die, $X = \#$ of heads

$E(X) = ?$ $\qquad X|Y \sim Bin(y, \frac{1}{2})$

$P_Y(y) = \frac{1}{6},\ y = 1, \dots, 6$

$$E(X) = E[\underbrace{E(X|Y=y)}] = E\left(y \cdot \frac{1}{2}\right) = \frac{1}{2}E(y)$$

$$= \frac{1}{2} \cdot 3.5 = 1.75 \ \searrow\ \frac{3+2.5+2+1.5+1+0.5}{6}$$

Theorem: $\boxed{Var(Y) = Var[E(Y|X)] + E[Var(Y|X)]}$

Proof: $Var(Y|X) = E(Y^2|X) - E(Y|X)^2$

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

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

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

we add and subtract $E[E(Y|X)^2]$

$= E(E(Y^2|X)) - E[E(Y|X)^2]$
$+ E[E(Y|X)^2] - (E[E(Y|X)])^2$

$= E[Var(Y|X)] + Var[E(Y|X)]$ ∎

Ex $X \sim Geom(p)$

$Y|X=x \sim Bin(x, p)$

$E(Y), Var(Y) = ?$

$E(Y) = E[E(Y|X)] = E(Xp) = pE(X)$

$= p \cdot \frac{1}{p} = 1$

$Var(Y) = Var(E(Y|X)) + E[Var(Y|X)] = Var(Xp) + E(Xpq)$
$= p^2 Var(X) + pq E(X) = \cancel{p^2} \cdot \frac{1-p}{\cancel{p^2}} + \cancel{p} q \frac{1}{\cancel{p}} = 2(1-p)$

Law of Large Numbers.

Toss a coin n times.

$$X_i = \begin{cases} 1, & \text{if } i^{th} \text{ toss is } H \\ 0, & \phantom{\text{if } i^{th} \text{ toss is }} T \end{cases}$$

$$X_i \sim \text{Bernoulli}\left(\tfrac{1}{2}\right), \quad E(X_i) = \frac{1}{2}$$

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \text{proportion of heads}$$

$$\overline{X}_n \xrightarrow[n\to\infty]{} \frac{1}{2}$$

Markov's Inequality: If X is a non-negative r.v. with $E(X) < \infty$ and $a > 0$ then

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Proof: (discrete case)

$$E(X) = \sum_x x\,p(x) = \underbrace{\sum_{x<a} x\,p(x)}_{\geq 0} + \sum_{x \geq a} x\,p(x)$$

$$E(X) \geq \sum_{x\geq a} x\,p(x) \geq \sum_{x \geq a} a\,p(x)$$

$$= a \sum_{x \geq a} p(x) = a\,P(X \geq a)$$

$$\Rightarrow \quad P(X \geq a) \leq \frac{E(X)}{a} \qquad \blacksquare$$

Chebyshev's Inequality: For a r.v. X with $E(X)<\infty$ and $Var(X)<\infty$, for any $a>0$

$$P(|X-E(X)|\geq a)\leq \frac{Var(X)}{a^2}$$

Proof: $Y=[X-E(X)]^2 \Rightarrow E(Y)=Var(X)$

$$P(|X-E(X)|\geq a)=P([X-E(X)]^2\geq a^2)$$

$$=P(Y\geq a^2)\leq \frac{E(Y)}{a^2}=\frac{Var(X)}{a^2}$$

by Markov's $\blacksquare$

Back to LLN:

Let $X_1, X_2, X_3, \ldots, X_n$ be independent and identically distributed (iid) r.v's. $E(X_i)=\mu$, $Var(X_i)=\sigma^2$

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$$

$$E(\overline{X}_n)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\mu$$

$$Var(\overline{X}_n)=\frac{1}{n^2}\sum_{i=1}^{n}Var(X_i)=\frac{\sigma^2}{n}$$

$$Var(\overline{X}_n)\xrightarrow[n\to\infty]{}0 \Rightarrow \overline{X}_n \xrightarrow[n\to\infty]{} \mu$$

Weak Law of Large Numbers (WLLN):

Suppose $X_1, X_2, \dots$ are iid with $E(X_i) = \mu < \infty$, $\underbrace{Var(X_i) = \sigma^2 < \infty}$, then $\forall a > 0$

$$\underbrace{P(|\bar{X}_n - \mu| \geq a) \to 0}, \text{ as } n \to \infty$$

$$\bar{X}_n \xrightarrow{P} \mu$$

($\bar{X}_n$ converges to μ in probability)

Proof: $E(\bar{X}_n) = \mu < \infty$

$$Var(\bar{X}_n) = \frac{\sigma^2}{n} \xrightarrow[n \to \infty]{} 0$$

$$P(|\bar{X}_n - \mu| \geq a) \leq \frac{Var(\bar{X}_n)}{a^2} = \frac{\sigma^2}{na^2} \xrightarrow[n \to \infty]{} 0$$

by Chebyshev's

$\blacksquare$

Ex. Flip a coin $\underbrace{10{,}000}$ times. Let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ toss is H} \\ 0, & \text{if } i^{th} \text{ toss is T} \end{cases}$

$$E(X_i) = \frac{1}{2}, \quad Var(X_i) = \frac{1}{4}$$

Take $a = 0.01$

$$P\left(\left|\bar{X}_n - \frac{1}{2}\right| \geq 0.01\right) \leq \frac{1}{4 \cdot 10{,}000 \cdot (0.01)^2} = \underbrace{\frac{1}{4}}$$

$$\frac{Var(\bar{X}_n)}{a^2}$$

weak upper bound

$$Var(\bar{X}_n) = \frac{1/4}{10{,}000}$$

Strong Law of Large Numbers (SLLN):

Suppose $X_1, X_2, X_3, \ldots$ are iid with $E(X_i) = \mu < \infty$,
then $\overline{X}_n$ converges to μ as $n \to \infty$ with probability 1:
(w.p. 1)

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \ldots + X_n}{n} = \mu\right) = 1$$

$\downarrow$
convergence almost surely