Lecture 10: Jan 28. Independent random variables Ross 6.2

X and Y are *independent* if for any subsets A and B of $\Re$, $P(X \in A \cap Y \in B) = P(X \in A) \times P((Y \in B)$.

10.1 Independence of two r.vs: discrete case

Discrete case: this is equivalent to $p_{X,Y}(x,y) = P(X=x, Y=y) = P(X=x).P(Y=y) = p_X(x)p_Y(y)$.

Clearly this is *necessary*: take $A = \{x\}$ and $B = \{y\}$.

Conversely, if $p_{X,Y}(x,y) = p_X(x)p_Y(y)$, then for any A, B:

$$\begin{aligned} P(X \in A,\ Y \in B) &= \sum_{x \in A}\sum_{y \in B} p_{X,Y}(x,y) = \sum_{x \in A}\sum_{y \in B} p_X(x)p_Y(y) \\ &= \sum_{x \in A} p_X(x) \sum_{y \in B} p_Y(y) = P(X \in A)\ P(Y \in B) \end{aligned}$$

Note this must hold for all x, y. Thus the ranges of the r.v.s cannot depend on each other. Example: X is value on first die, Y on second: these are independent. But if we throw out doubles (i.e. points with $x = y$) they are no longer independent: if $X = 4$ we know $Y \neq 4$.

10.2 Independence of two r.vs: continuous case

With $A = (-\infty, x)$ and $B = (-\infty, y)$ we see $F_{X,Y}(x,y) = F_X(x)F_Y(y)$.

As with the 1-dimensional case, this is also sufficient:

X and Y are independent if and only if $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ for all x, y.

Differentiating, we see this means $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, and conversely integrating we see these are equivalent. Note again it must hold for all x, y: the ranges of the r.vs cannot depend on each other.

10.3 Buffon's needle: a classic example of estimating π.

Parallel lines distance D apart; needle length L, with $L \leq D$.

Let X be distance from needle midpoint to nearest line; θ be angle of needle to X. X is $U(0, D/2)$, θ is $U(0, \pi/2)$ and they are *independent*. We want the probability

$$\begin{aligned} P(X \leq (L/2)\cos(\theta)) &= \int\int_{2x < L\cos(y)} f_X(x) f_\theta(y)\ dx\ dy = \frac{2}{\pi}\frac{2}{D}\int_0^{\pi/2}\int_0^{L/2\cos(y)} dx\ dy \\ &= \frac{4}{\pi\ D}\int_0^{\pi/2}\frac{L}{2}\cos(y)\ dy = \frac{2\ L}{\pi\ D}[\sin(y)]_0^{\pi/2} = 2L/\pi D \end{aligned}$$

10.4 Examples of independence and dependence

X and Y are independent if and only if $f_{X,Y} = f_X(x)f_Y(y)$ for all x, y.

Also, if $f_{X,Y} = g_1(x)g_2(y)$ for all X, Y, then X and Y are independent, and $f_X(x) \propto g_1(x)$, $f_Y(y) \propto g_2(y)$.

(i) Example: $f_{X,Y}(x,y) = x + y$ on $0 < x < 1$, $0 < y < 1$. (Ch 6: # 22)

X and Y are not independent: why not?

(ii) Example: $f_{X,Y}(x,y) = \exp(-(x+y))$ on $0 < x < \infty$, $0 < y < \infty$.

X and Y are independent: why? What is $f_X(x)$?

(iii) Example: $f_{X,Y}(x,y) = 2\exp(-(4x + \frac{1}{2}y))$ on $0 < x < \infty$, $0 < y < \infty$.

X and Y are independent: why? What is $f_X(x)$?

(iv) Example: $f_{X,Y}(x,y) = 1/6$ on $0 < x < 2$, $0 < y < 3$,

X and Y are independent: why? What is $f_X(x)$?

(v) Example: $f_{X,Y}(x,y) = 2$ on $0 < x$, $0 < y$, $x + y < 1$

X and Y are not independent: why not?

Lecture 11: Jan 30. Examples of computing probabilities

11.1 Computing marginal densities

(i) Example: $f(x,y) = x + y$ on $0 < x < 1$, $0 < y < 1$. (Ch 6: # 22)

What is $f_X(x)$?

(ii) Example: $f_{X,Y}(x,y) = 2$ on $0 < x$, $0 < y$, $x + y < 1$

X and Y are not independent: why not?

What is $f_X(x)$?

(ii) Example: $f_{X,Y}(x,y) = 2\exp(-(4x + \frac{1}{2}y))$ on $0 < x < \infty$, $0 < y < \infty$.

X and Y are independent: why? What is $f_X(x)$?

How would you compute $f_X(x)$ if you did not see X and Y are independent?

11.2 Computing probabilities with joint pdf's

Recall: $P((X,Y) \in A) \;=\; \int\int_A f_{X,Y}(x,y)\; dx\; dy.$

(i) Example: $f(x,y) = x + y$ on $0 < x < 1$, $0 < y < 1$. (Ch 6: # 22)

$$\begin{aligned} P(X+Y<1) &= \int_{y=0}^{1}\int_{x=0}^{1-y}(x+y)\; dx\; dy \;=\; \int_{y=0}^{1}[\frac{1}{2}x^2 + yx]_0^{1-y}dy \\ &= \int_{y=0}^{1}(\frac{1}{2}(1-y)^2 + y(1-y))dy \;=\; \int_{y=0}^{1}\frac{1}{2}(1-y^2)dy \;=\; \frac{1}{2}(1-(1/3)) \;=\; 1/3. \end{aligned}$$

(ii) Example: $f(x,y) = 2$ on $0 < x$, $0 < y$, $x + y < 1$

$$\begin{aligned} P(X<3Y) &= \int_{y=0}^{1}\int_{x=0}^{\min(1-y,3y)} 2\; dx\; dy \;=\; \int_{y=0}^{1/4} 6y\; dy \;+\; \int_{y=1/4}^{1} 2(1-y)\; dy \\ &= [3y^2]_0^{1/4} \;+\; [-(1-y)^2]_{1/4}^{1} \;=\; 3/16 \;+\; 9/16 \;=\; 3/4. \end{aligned}$$

11.3: Sum of two independent U(0,1)

Let $X \sim U(0,1)$ and $Y \sim U(0,1)$ with X and Y independent.

So $f_{X,Y}(x,y) \;=\; 1$ on $0 < x < 1$, $0 < y < 1$.

(a) What is the range of $W = X + Y$?

(b) Compute $P(X + Y < w)$ where $0 < w < 1$.

Or, draw a picture. (Answer: $w^2/2$)

(c) Hence show $f_W(w) = w$ if $0 < w < 1$. (This is only part of $f_W(w)$.)

(c) Note $(2 - W) \;=\; (1 - X) \;+\; (1 - Y)$.

Why does this tell me that $f_W(w)$ is symmetric about $w = 1$?

(d) Because of its shape, $f_W(w)$ is known as a triangular density.

Do these densities form a location/scale family ?

Lecture 12: Feb 2: Conditional pmf (discrete) and pdf (continuous): two examples

12.1 Convolution of probability mass functions

Let X and Y be independent discrete r.vs with pmf $p_X()$ and $p_Y()$.

$$P(W \equiv (X+Y) = w) \quad = \quad \sum_x P(W = w \cap X = x) \quad = \quad \sum_x P(Y = w-x \cap X = x) \quad = \quad \sum_x p_Y(w-x) p_X(x)$$

Example: sum of independent Poissons; recall the Poisson process number of events.

Let X be $\mathcal{P}o(\mu)$ and Y be $\mathcal{P}o(\nu)$, and X,Y independent.

$$\begin{aligned} P(X+Y=k) &= \sum_{x=0}^{k} P(Y=(k-x))P(X=x) \quad = \quad \sum_{x=0}^{k} \frac{\exp(-\nu)\nu^{k-x}}{(k-x)!} \frac{\exp(-\mu)\mu^x}{x!} \\ &= \frac{\exp(-(\mu+\nu))}{k!} \sum_{x=0}^{k} \frac{k!}{(k-x)!x!} \mu^x \ \nu^{k-x} \quad = \quad \frac{\exp(-(\mu+\nu))}{k!} (\mu+\nu)^k \end{aligned}$$

(Remember the binomial theorem.)

12.2 Conditional pmf (Ross 6.4): $P(X = x | W = w) \ = \ p_{X,W}(x,w)/p_W(w)$

Example: Let X be $\mathcal{P}o(\mu)$ and Y be $\mathcal{P}o(\nu)$, and X,Y independent, and $W = X+Y$. We know $W \ \sim \ \mathcal{P}o(\mu+\nu)$.

$$\begin{aligned} P(X = x \mid W = w) &= P(X=x, Y=w-x)/P(W=w) \ = \ p_X(x)p_Y(w-x)/P(W=w) \\ &= \frac{\exp(-\mu)\mu^x}{x!} \frac{\exp(-\nu)\nu^{w-x}}{(w-x)!} \frac{w!}{\exp(-(\mu+\nu))(\mu+\nu)^w} \ = \ \binom{w}{x} \left(\frac{\mu}{\mu+\nu}\right)^x \left(\frac{\nu}{\mu+\nu}\right)^{w-x} \end{aligned}$$

i.e. $X \mid W = w$ is Binomial $(w, \ \mu/(\mu+nu))$.

12.3 Convolution of a probability density function (Ross 6.3)

X and Y are independent, with $f_{X,Y}(x,y) = f_X(x) f_Y(y)$.

$$\begin{aligned} F_{X+Y}(a) &= \int\int_{x+ylea} f_X(x) \ f_Y(y) \ dx \ dy \ = \ \int_{y=-\infty}^{\infty} f_Y(y) \left(\int_{x=-\infty}^{a-y} f_X(x) \ dx \right) \ dy \ = \ \int_{-\infty}^{\infty} F_X(a-y) \ f_Y(y) \ dy \\ f_{X+Y}(a) &= \frac{d}{da}\left(\int_{-\infty}^{\infty} F_X(a-y) \ f_Y(y) \ dy\right) \ = \ \int_{-\infty}^{\infty} f_x(a-y) \ f_Y(y) \ dy \end{aligned}$$

We could use this to show that if $X \ \sim \ G(m,\lambda)$, $Y \ \sim \ G(n,\lambda)$ then $X+Y \ \sim \ G(m+n,\lambda)$ (see Ross P.281), but we know this anyhow from the Poisson process, times to m^{th} then n^{th} events.

12.4 Conditional probability density function: define $f_{X|Y}(x|y) \ = \ f_{X,Y}(x,y)/f_Y(y)$**.** (Ross 6.5)

$$P(x < X \le x+\delta x \mid y < Y \le y+\delta y) \quad = \quad \frac{P(x < X \le x+\delta x \cap y < Y \le +\delta y)}{P(y < Y \le y+\delta y)} \ \approx \ \frac{f_{X,Y}(x,y) \ \delta x \ \delta y}{f_Y(y) \ \delta y}$$

Example: $X \ \sim \ G(m,\lambda)$, $Y \ \sim \ G(n,\lambda)$, independent, so $W \equiv X+Y \ \sim \ G(m+n,\lambda)$

$$\begin{aligned} f_{X|W}(x|w) &= f_{X,W}(x,w)/f_W(w) \ = \ f_X(x) \ f_Y(w-x)/f_W(w) \\ &= \frac{\lambda^m x^{m-1}\exp(-\lambda x)}{\Gamma(m)} \frac{\lambda^n (w-x)^{n-1}\exp(-\lambda(w-x))}{\Gamma(n)} \frac{\Gamma(m+n)}{\lambda^{m+n} w^{m+n-1}\exp(-\lambda w)} \\ &= \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} w^{-1}(x/w)^{m-1}(1-(x/w))^{n-1} \quad \text{on } 0 \le x \le w \quad \text{and 0 otherwise} \end{aligned}$$

Note λ is gone, but w is a scale parameter. In fact, if $V = X/W$,

$f_{V|W}(v|w) \ = \ \Gamma(m+n)v^{m-1}(1-v))^{n-1}/\Gamma(m)\Gamma(n)$ on $0 \le v \le 1$, which does not depend on w.