Lecture 15: Fri Oct 26 (a): Joint densities L & M 3.7

15.1 Joint and marginal (cumulative) distribution functions

For two random variables X and Y the *joint cdf* is $F_{X,Y}(a,b) \; = \; P(X \leq a, \; Y \leq b), \;\; -\infty < a, b < \infty$.
Note that the *marginal cdfs* of X and of Y are given by

$$\begin{aligned} F_X(a) \;&=\; P(X \leq a) \;=\; P(\lim_{b\to\infty} \{X \leq a, Y \leq b\}) \quad \text{(increasing sequence of sets)} \\ &=\; \lim_{b\to\infty} P(X \leq a, \; Y \leq b) \;=\; \lim_{b\to\infty} F_{X,Y}(a,b) \;\equiv\; F_{X,Y}(a,\infty) \\ \text{Similarly,} \quad F_Y(b) \;&=\; P(Y \leq b) \;=\; \lim_{a\to\infty} F_{X,Y}(a,b) \;\equiv\; F_{X,Y}(\infty, b) \end{aligned}$$

Just as in 1 dimension, we can get all other probabilities from $F_{X,Y}$. For example (draw picture):

$$P(a_1 < X \leq a_2, \; b_1 < Y \leq b_2) \;=\; F_{X,Y}(a_2,b_2) - F_{X,Y}(a_1,b_2) \;-\; F_{X,Y}(a_2,b_1) \;+\; F_{X,Y}(a_1,b_1)$$

15.2 Joint and marginal probability mass functions

If X and Y are discrete random variables the *joint pmf* is $p_{X,Y}(x,y) \;=\; P(X=x, y=y) \quad$ for $x \in \mathcal{X}, \; y \in \mathcal{Y}$.
Then the *marginal pmfs* of X and of Y are

$$p_X(x) \;=\; P(X=x) \;=\; \sum_{y\in\mathcal{Y}} p_{X,Y}(x,y), \qquad \text{and } p_Y(y) \;=\; \sum_{x\in\mathcal{X}} p_{X,Y}(x,y).$$

Note $p_X(x) > 0$ for $x \in \mathcal{X}$, and $p_Y(y) > 0$ for $y \in \mathcal{Y}$, but $p_{X,Y}(x,y)$ can be 0 for some $x \in \mathcal{X}, \; y \in \mathcal{Y}$.

15.3 The multinomial distribution

If we take a sample size n (with replacement) and there are r types, with type i having probability p_i, $i = 1, ..., r$ and $\sum_{i=1}^r p_i = 1$. Let X_i be the number of type i in the sample ($i = 1, ..., r$): $\sum_{i=1}^r X_i = n$:

$$P(X_1 = n_1, \; X_2 = n_2, ..., X_r = n_r) \;=\; \frac{n!}{n_1! \; n_2! .... n_k!} p_1^{n_1} p_2^{n_2} .... p_r^{n_r}$$

Recall (Lecture 3), number of ways of arranging n_1 objects type-1, n_2 objects type-2, ... n_k objects type-k, where $n_1 + n_2 + .. + n_k \;=\; n$ is $n!/(n_1! \; n_2! .... n_k!)$.

Example: There are 4 ABO blood types, A, B, AB and O. For the USA population, roughly, $P(A) = 0.41$, $P(B) = 0.16$, $P(AB) = 0.07$, and $P(O) = 0.36$. Twelve students go to donate blood: what is the probability 5 are type A, 2 are type B, one is AB, and 4 are type O? Answer:

$$12!/(5! \times 2! \times 1! \times 4!)(0.41)^5(0.16)^2(0.07)^1(0.36)^4 \;=\; 914760 \times 3.487^{-7} \;\approx\; 0.319$$

What is the (marginal) pmf of the number of students with type A blood? Answer: Bin(12, 0.41). (Why?)

15.4 Joint and marginal probability density functions

(i) Random variables X and Y are *jointly continuous* if there is a function $f_{X,Y}(x,y)$ defined for all real x and y, such that for every (? not quite?) set C in $\Re^2$, $P((X,Y) \in C) \;=\; \int\int_{(x,y)\in C} f_{X,Y}(x,y) \; dx \; dy$.
Then $f_{X,Y}(x,y)$ is the *joint pdf* of X and Y.

(ii)
$$F_{X,Y}(a,b) \;=\; P(X \in (-\infty, a], \; Y \in (-\infty, b]) \;=\; \int_{y=-\infty}^{b} \int_{x=-\infty}^{a} f_{X,Y}(x,y) \; dx \; dy$$

$$\text{so } f_{X,Y}(a,b) \;=\; \frac{\partial^2}{\partial a \; \partial b} F_{X,Y}(a,b)$$

(iii)
$$P(X \in A) \;=\; P(X \in A, \; Y \in (-\infty,\infty)) \;=\; \int_{X\in A} \int_{y=-\infty}^{\infty} f_{X,Y}(x,y) \; dx \; dy \;=\; \int_{X\in A} f_X(x) \; dx$$

$$\text{where } f_X(x) \;=\; \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$$

So $f_X(x)$ is pdf of X and similarly pdf of Y is $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$

Lecture 16: Oct 26 (b). Independent random variables (L & M 3.7,3.8)

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

16.1 Independence of two r.vs: discrete case

Theorem: Independence is equivalent to $p_{X,Y}(x,y) = P(X = x, Y = y) \ = \ P(X = x).P(Y = y) = p_X(x)p_Y(y)$.

Proof: 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 $x \in \mathcal{X}, \ y \in \mathcal{Y}$. Thus independence implies $p_{X,Y}(x,y) > 0$ for all $x \in \mathcal{X}, \ yi \in \mathcal{Y}$.

If X and Y are independent, 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 $Y \neq 4$.

16.2 Independence of two r.vs: continuous case

With $A = (-\infty, x)$ and $B = (-\infty, y)$ we see

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

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

Conversely: If $f_{X,Y}(x,y) \ = \ f_X(x)f_Y(y)$ then

$$P(X \in A, Y \in B) \ = \ \int_{XinA} \int_{Y \in B} f_{X.Y}(x,y)dxdy \ = \ \int_{XinA} \int_{Y \in B} f_X(x)f_Y(y)dxdy \ = \ P(X \in A)P(Y \in B)$$

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

Integrating f back to F: X and Y are independent if and only if $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ 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 $f_X(x) \propto g_1(x)$, $f_Y(y) \propto g_2(y)$.

Note again these must hold for all $x, \ y$: the ranges of the r.vs cannot depend on each other.

16.3 Examples of independence and dependence

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

X and Y are not independent: why not? What is $f_X(x)$?

(ii) Example: $f_{X,Y}(x,y) = 24xy$ on $0 < x$, $0 < y$, $x + y < 1$

X and Y are not independent: why not? What is $f_X(x)$?

(Compare this with $f_{X,Y}(x,y) = 4xy$ on $0 < x < 1$, $0 < y < 1$, where X and Y are independent.)

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

or $f_{X,Y}(x,y) \ = \ \exp(-(x + y))I_{(0,\infty)}(x)I_{(0,\infty)}(y)$

X and Y are independent: why? What is $f_X(x)$? (Note: $I_A(x) = 1$ if $x \in A$ and 0 oterwise.)

(iv) 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?

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

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

17: Examples of computing probabilities

17.1 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 $0<x$, $0<y$, $x+y<1$

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

Easier: Draw a picture. For a constant density. Probability $\equiv$ area.

17.2: 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 $w=1$?

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

17.3 Buffon's needle: a classic example of estimating π. (For interest only)

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

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

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

17.4 The variance of a Normal N(0,1): For interest only.

A $N(0,1)$ r.v. had p.d.ff $f_Z(z) \;=\; \frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2})$ on $-\infty<z<\infty$ is a *standard Normal random variable.*

$$\begin{aligned} \left(\int_{-\infty}^{\infty}\exp(-z^2/2)dz\right)^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-(x^2+y^2)/2)dx\;dy \;=\; \int_0^{\infty}\int_0^{2\pi}\exp(-r^2/2)\;r\;dr\;d\theta \\ &= [-\exp(-r^2/2)]_0^{\infty}\;[\theta]_0^{2\pi} = 2\pi \qquad \text{using polar coordinates} \end{aligned}$$

$$\mathrm{E}(Z) \;=\; \int_{-\infty}^{\infty} z f_Z(z)dz \;=\; 0 \qquad \text{since } f_Z(-z) = f_Z(z).$$

$$\begin{aligned} \mathrm{var}(Z) \;=\; \mathrm{E}(Z^2) \;=\; \int_{-\infty}^{\infty} z^2 f_Z(z)dz &= \frac{1}{\sqrt{2\pi}}\left([-z.exp(-z^2/2)]_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\exp(-z^2/2)dz\right) \\ &= \int_{-\infty}^{\infty} z f_Z(z)dz \;=\; 1 \end{aligned}$$

Lecture 18: Transforming random variables Wed Oct 31 L & M 3.8

18.1 Linear transformations: Location and scale

Let X have pdf $f_X(x)$, $\mathrm{E}(X) = \mu$, $\mathrm{var}(X) = \sigma^2$, and $Y = aX + b$, $X = (Y - b)/a$.

(a) **Location:** $(a = 1)$: $Y = X + b$, $\mathrm{E}(Y) = \mu + b$, $\mathrm{var}(Y) = \sigma^2$:
$F_Y(y) \;=\; P(Y \leq y) \;=\; P(X \leq y - b) = F_X(y - b)$. So $f_Y(y) = f_X(y - b)$.

(b) **Scale:** $(b = 0,\ a > 0)$: $Y = aX$, $\mathrm{E}(Y) = a\mu$, $\mathrm{var}(Y) = a^2\sigma^2$:
$F_Y(y) \;=\; P(Y \leq y) \;=\; P(X \leq y/a) \;=\; F_X(y/a)$. So $f_Y(y) \;=\; (1/a) f_X(y/a)$.

(c) **Location and scale:** $(a > 0)$: $Y = aX + b$, $\mathrm{E}(Y) = a\mu + b$, $\mathrm{var}(Y) = a^2\sigma^2$:
$F_Y(y) \;=\; P(Y \leq y) \;=\; P(X \leq (y - b)/a) \;=\; F_X((y - b)/a)$. So $f_Y(y) \;=\; (1/a) f_X((y - b)/a)$.

18.2 Exponential random variables have scale only

Note that exponential densities have scale λ^{-1}, where λ is the rate parameter.

Instead of a density $(1/a)f(y/a)$ with scale parameter a, we have a density of form $\lambda f(\lambda y)$. (i.e. $a = (1/\lambda)$)

If $Y \;\sim\; \mathcal{E}(\lambda)$, then $\lambda Y \;\sim\; \mathcal{E}(1)$.

If buses come at rate 0.1 per minute, they come at rate $60 \times 0.1 \;= 6$ per hour.

My expected waiting time is 10 minutes, or $10/60 \;=\; 1/6$ hours.

18.3 Normal random variables have location and scale

The Normal r.v. $X \;\sim\; N(\mu, \sigma^2)$ has pdf $f_X(x) = (1/\sqrt{2\pi})(1/\sigma)\exp(-(1/2)((x - \mu)/\sigma)^2)$.

So $f_X(x) \;=\; (1/\sigma) f_Z((x - \mu)/\sigma)$ where $f_Z(z) \;=\; (1/\sqrt{2\pi})\exp(-(1/2)z^2)$ is pdf of $N(0, 1)$.

Recall, $X \;\sim\; N(\mu, \sigma^2)$ gives $Z = (X - \mu)/\sigma \;\sim\; N(0, 1)$.

Conversely, $Z \;\sim N(0, 1) \;\;\Rightarrow\;\; X = \mu + \sigma Z \;\sim\; N(\mu, \sigma^2)$.

If we assume $\mathrm{E}(Z) = 0$ and $\mathrm{var}(Z) = 1$ (we haven't shown it yet), then $\mathrm{E}(X) = \mu$ and $\mathrm{var}(X) = \sigma^2$.

That is, the two parameters that define the Normal pdf are the mean and the variance.

Now let $Y \;=\; aX + b \;=\; a(\mu + \sigma Z) + b \;=\; (a\mu + b) + a\sigma Z$. So $Y \;\sim\; N(a\mu + b, a^2\sigma^2)$.

18.4 Uniform random variables have location and scale

Uniform random variables, $U(a, b)$ have both location and scale: linear functions of uniforms are uniform.

Let X be $U(0, 1)$: $f_X(u) \;=\; 1$ for $0 \leq u \leq 1$ and 0 otherwise. Or $f_X(x) = I_{(0,1)}(x)$.

Note $\mathrm{E}(X) = 1/2$, $\mathrm{var}(X) = 1/12$.

Note $F_X(u) = 0$ for $u \leq 0$, $F_X(u) = u$ for $0 \leq u \leq 1$, and $F_X(u) = 1$ for $u \geq 1$.

Now let $Y \;=\; c + (k - c)X$ where $c < k$: $0 \leq X \leq 1$ so $c \leq Y \leq k$. Consider $c \leq y \leq k$.

Then $F_Y(y) \;=\; P(Y \leq y) \;=\; P(c + (k - c)X \leq y) \;=\; P(X \leq (y - c)/(k - c)) \;=\; (y - c)/(k - c)$.

So $f_Y(y) = 1/(k - c)$, if $c \leq y \leq k$ and 0 otherwise.

That is

$$f_Y(y) \;=\; \frac{1}{(k - c)} I_{(c,k)}(y) \;=\; \frac{1}{k - c} I_{(0,1)}\left(\frac{y - c}{k - c}\right)$$

So we have a location-scale family with location c and scale $(k - c)$.

18.5 Example: Do the triangular densities of **17.2** form a location/scale family?