Lecture 32: Nov. 14

Last time

- Presentations
- Multiple Random Variables (Chapter 4)

Today

- Conditional Distributions

Example 2

$$
\begin{aligned}
F_{X,Y}(x,y) &= x - x\log\frac{x}{y} \qquad 0 < x \leqslant y \leqslant 1 \\
f_{X,Y}(x,y) &= \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \\
f_X(x) &= \\
f_Y(y) &=
\end{aligned}
$$

Note: Once we have $f_X(x)$ and $f_Y(y)$, we can obtain $F_X(x)$ and $F_Y(y)$ directly. Double check: $F_X(x) = F_{X,Y}(x,\infty)$.

Conditional Distributions

Conditional Distributions - Discrete Recall if A and B are two events, the probability of A conditional on B is:

$$\Pr(A|B) = \frac{\Pr(A,B)}{\Pr(B)}$$

Defining the events $A = \{Y = y\}$ and $B = \{X = x\}$, it follows that

$$
\begin{aligned}
\Pr\{Y = y|X = x\} &= \frac{\Pr(X = x, Y = y)}{\Pr(X = x)} \\
&= \frac{f_{X,Y}(x,y)}{f_X(x)} \\
&= f_{Y|X}(y|x)
\end{aligned}
$$

This is called the *conditional probability mass function* of Y given X.

Example: Discrete Back to the fair coin example. From the joint pmf of X and Y, we can derive all the conditional pmfs:

		Y			
		0	1	2	3
X	0	1/8	1/4	1/8	0
	1	0	1/8	1/4	1/8

Conditional Distribution - Continuous If $F(x, y)$ is absolutely continuous, we define the conditional density of Y given X as:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \text{ if } f_Y(y) > 0$$

Example 1

$$\begin{array}{ll}
F_{XY}(x,y) = xy & 0 < x < 1,\ \ 0 < y < 1 \\
f_{XY}(x,y) = 1 & 0 < x < 1,\ \ 0 < y < 1 \\
f_X(x) = 1 & 0 < x < 1 \\
f_Y(y) = 1 & 0 < y < 1 \\
f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = 1 & 0 < x < 1 \quad (0 < y < 1) \\
f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = 1 & 0 < y < 1 \quad (0 < x < 1)
\end{array}$$

Note: Here we get that the conditional densities are the same as the marginals. This means X and Y are independent.

Example 2

$$\begin{array}{lll}
F_{XY}(x,y) = x - x\log\frac{x}{y} & 0 < x \leqslant y \leqslant 1 & \\
f_{XY}(x,y) = 1/y & 0 < x \leqslant y \leqslant 1 & \\
f_X(x) = -\log x & 0 < x \leqslant 1 & \\
f_Y(y) = 1 & 0 < y \leqslant 1 & \\
f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = 1/y & 0 < x \leqslant y & (0 < y \leqslant 1) \\
f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = -\frac{1}{y\log x} & x \leqslant y \leqslant 1 & (0 < x \leqslant 1)
\end{array}$$

- Y is marginally uniform, but not conditionally uniform.
- X is conditionally uniform, but not marginally uniform.

Independent Random Variables

Independence The random variable X and Y are said to be *independent* if for any two Borel sets A and B,

$$\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B)$$

All events defined in terms of X are independent of all events defined in terms of Y.

Using the Kolmogorov axioms of probability, it can be shown that X and Y are independent if and only if $\forall(x, y)$ (except possibly for sets of probability 0)

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

or in terms of pmfs (discrete) and pdfs (continuous)

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

Checking independence

- A necessary condition for independence of X and Y is that their joint pdf/pmf has positive probability on a rectangular domain.
- If the domain is rectangular, one can try to write the joint pdf/pmf as a product of functions of x and y only.

Example Two points are selected randomly on a line of length a so as to be on opposite sides of the mid-point of the line. Find the probability that the distance between them is less than $a/3$.

Solution:
Let X be the coordinate of a point selected randomly in $[0, a/2]$ and Y be the coordinate of a point selected randomly in $[a/2, a]$. Assume X and Y are independent and uniform over its interval. The joint density is

$$f_{X,Y}(x, y) = 4/a^2, \quad 0 \leqslant x \leqslant a/2, a/2 \leqslant y \leqslant a$$

Therefore, the solution is

$$\Pr(Y - X < a/3) =$$

Example: Buffon's Needle A table is ruled with lines distance 1 unit apart. A needle of length $L \leqslant 1$ is thrown randomly on the table. What is the probability that the needle intersects a line?
Solution:
Define two random variables:

- X: distance from low end of the needle to the nearest line above
- θ: angle from the vertical to the needle.

By "random", we assume X and θ are independent, and

$$X \sim U(0, 1) \quad \text{and} \quad \theta \sim U[-\pi/2, \pi/2].$$

This means that

$$f_{X,\theta}(x, \theta) = 1/\pi, \quad 0 \leqslant x \leqslant 1, -\pi/2 \leqslant \theta \leqslant \pi/2$$

For the needle to intersect a line, we need $X < L\cos(\theta)$.

Expectations of Independent RVs (Theorem 4.2.10) Let X and Y be independent rvs.

- For any $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$,

$$\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B)$$

  i.e. the events $\{X \in A\}$ and $\{Y \in B\}$ are independent.

- Let $g(x)$ be a function only of x and $h(y)$ be a function only of y. Then

$$E\left[g(X)h(Y)\right] = \left[Eg(X)\right]\left[Eh(Y)\right]$$

Proof:

$$\begin{aligned}
E\left[g(X)h(Y)\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_{XY}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\
&= \left(\int_{-\infty}^{\infty} g(x)f_X(x)dx\right)\left(\int_{-\infty}^{\infty} h(y)f_Y(y)dy\right) \\
&= \left[Eg(X)\right]\left[Eh(Y)\right]
\end{aligned}$$

Example X, Y are independent

$$\begin{aligned}
E(X^2Y^3) &= (EX^2)(EY^3) \\
E(Y^2Y^3) &\neq (EY^2)(EY^3)
\end{aligned}$$

Bivariate Transformation

Functions of random variables Let (X,Y) be a bivariate rv with known distributions. Define (U,V) by

$$U = g_1(X,Y), \quad V = g_2(X,Y)$$

Probability mapping For any Borel set $B \subset \mathbb{R}^2$,

$$\Pr\left[(U,V) \in B\right] = \Pr\left[(X,Y) \in A\right]$$

where A is the inverse mapping of B, i.e.

$$A = \{(x,y) \in \mathbb{R}^2 : (g_1(x,y), g_2(x,y)) \in B\}$$

The inverse is well defined even if the mapping is not bijective.

Example Let $g_1(x,y) = x, g_2(x,y) = x^2 + y^2$.

Discrete RVs Suppose that (X,Y) is a discrete rv, i.e. the pmf is positive on a countable set $\mathcal{A}$. Then (U,V) is also discrete and takes values on a countable set $\mathcal{B}$. Define

$$A_{u,v} = \{(x,y) \in \mathcal{A} : g_1(x,y) = u, g_2(x,y) = v\}$$

Then

$$f_{UV}(u,v) = \Pr(U=u, V=v) = \sum_{(x,y)\in A_{u,v}} f_{XY}(x,y)$$

Sum of two independent Poissons Let $X \sim Poisson(\lambda_1)$, $Y \sim Poisson(\lambda_2)$, independent, and define

$$U = X + Y, \quad V = Y$$

- (X, Y) takes values in $\mathcal{A} = \{0, 1, 2, \dots\}^2$
- (U, V) takes values on $\mathcal{B} = \{(u, v) : v = 0, 1, 2, \dots, u = v, v + 1, v + 2, \dots\}$.
- For a particular (u, v), $A_{uv} = \{(x, y) \in \mathcal{A} : x + y = u, y = v\} = (u - v, u)$.

The joint pmf of U and V is

$$f_{UV}(u, v) = f_{XY}(u - v, v) = \frac{e^{-\lambda_1}\lambda_1^{u-v}}{(u-v)!}\frac{e^{-\lambda_2}\lambda_2^{v}}{(v)!}$$

The distribution of $U = X + Y$ is the marginal

$$\begin{aligned} f + U(u) &= \sum_{v=0}^{u} \frac{e^{-\lambda_1}\lambda_1^{u-v}}{(u-v)!}\frac{e^{-\lambda_2}\lambda_2^{v}}{(v)!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{u!}\sum_{v=0}^{u}\binom{u}{v}\lambda_1^{u-v}\lambda_2^{v} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{u!}(\lambda_1+\lambda_2)^u \end{aligned}$$

We obtain that U is Poisson with parameter $\lambda = \lambda_1 + \lambda_2$.

Bivariate Transformations of Continuous RVs Suppose (X, Y) is continuous and the joint transformation

$$u = g_1(x, y), \quad v = g_2(x, y)$$

is one-to-one and differentiable. Define the inverse mapping

$$x = h_1(u, v), \quad y = h_2(u, v)$$

Then

$$f_{UV}(u, v) = f_{XY}\left(h_1(u, v), h_2(u, v)\right)|J(u, v)|$$

where $J(u, v)$ is the Jacobian of the transformation $(x, y) \to (u, v)$ given by

$$J(u, v) = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

Example: Rotation of a bivariate normal vector Let $X \sim N(0, 1)$, $Y \sim N(0, 1)$, independent. Define the rotation

$$\begin{aligned} U &= X\cos\theta - Y\sin\theta \\ V &= X\sin\theta + Y\cos\theta \end{aligned}$$

for fixed θ. Then $U \sim N(0, 1)$, $V \sim N(0, 1)$, independent.

Proof:
The range of (X, Y) is $\mathbb{R}^2$. The range of (U, V) is $\mathbb{R}^2$. Need the inverse transformation

$$\begin{aligned} X &= U\cos\theta + V\sin\theta \\ Y &= -U\sin\theta + V\cos\theta \end{aligned}$$

with Jacobian

$$J(u, v) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{vmatrix} = 1$$

The joint pdf of (X, Y) is

$$f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

The joint pdf of (U, V) is

$$\begin{aligned} f_{UV}(u, v) &= \frac{1}{2\pi} e^{-\left[(u\cos\theta + v\sin\theta)^2 + (-u\sin\theta + v\cos\theta)^2\right]/2} \cdot |1| \\ &= \frac{1}{2\pi} e^{-(u^2+v^2)/2} = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-v^2/2} \end{aligned}$$

so $U \sim N(0, 1)$, $V \sim N(0, 1)$, and U and V are independent.

Functions of independent random variables (Theorem 4.3.5) Let X and Y be independent rvs. Let $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ be functions. Then the random variables $U = g(X)$ and $V = h(Y)$ are independent.

Sum of two independent rvs Suppose X and Y are independent. What is the distribution of $Z = X + Y$? In general:

$$F_Z(z) = \Pr(X + Y \leqslant z) = \Pr(\{(x, y) \text{ such that } x + y \leqslant z\})$$

Various approaches:

- bivariate transformation method (continuous and discrete)
- Discrete convolution

$$f_Z(z) = \sum_{x+y=z} f_X(x) f_Y(y) = \sum_x f_X(x) f_Y(z - x)$$

- Continuous convolution (Section 5.2)
- MGF method (continuous and discrete)

Example (Sum of two independent Poissons) Define X, Y to be two independent random variables having Poisson distributions with parameters λ_i, $i = 1, 2$. Then:

$$f_{X,Y}(x,y) = \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2}\lambda_2^y}{y!}, x, y = 0, 1, 2, \ldots$$

The distribution of $S = X + Y$ is

$$\begin{aligned} f_S(s) &= \sum_{x=0}^{s} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2}\lambda_2^{s-x}}{(s-x)!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{s!}\sum_{x=0}^{s}\binom{s}{x}\lambda_1^x\lambda_2^{s-x} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{s!}(\lambda_1+\lambda_2)^s \end{aligned}$$

Again, S is Poisson with parameter $\lambda = \lambda_1 + \lambda_2$.

Moment generating function (Theorem 4.2.12) Let X and Y be independent rvs with mgfs $M_X(\cdot)$ and $M_Y(\cdot)$, respectively. Then the mgf of $Z = X + Y$ is

$$M_Z(t) = M_X(t)M_Y(t)$$

Proof:

$$\begin{aligned} M_Z(t) &= E\exp(Zt) & &= E\{\exp[(X+Y)t]\} \\ &= E[\exp(Xt)\exp(Yt)] & &= E[\exp(Xt)] \cdot E[\exp(Yt)] \\ &= M_X(t)M_Y(t) \end{aligned}$$

Corollary: If X and Y are independent and $Z = X - Y$,

$$M_Z(t) = M_X(t)M_Y(-t)$$

Example (sum of two independent Poissons) Suppose $X \sim Poisson(\lambda_X)$ and $Y \sim Poisson(\lambda_Y)$ and put $Z = X + Y$. Then, $Z \sim Poisson(\lambda_X + \lambda_Y)$. *Proof:*

$$\begin{aligned} M_Z(t) &= \exp\left[\lambda_X(e^t-1)\right]\exp\left[\lambda_Y(e^t-1)\right] \\ &= \exp\left[(\lambda_X+\lambda_Y)(e^t-1)\right] \end{aligned}$$

Example (sum of two independent normals) Suppose $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$ and X and Y are independent and $Z = X + Y$. Then

$$Z \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

Proof:

$$\begin{aligned} M_Z(t) &= \exp\left(\mu_x t + \frac{1}{2}\sigma_x^2 t^2\right)\exp\left(\mu_y t + \frac{1}{2}\sigma_y^2 t^2\right) \\ &= \exp\left[(\mu_x+\mu_y)t + \frac{1}{2}(\sigma_x^2+\sigma_y^2)t^2\right] \end{aligned}$$

Example (sum of two independent gammas) Suppose $X \sim \Gamma(\alpha_x, \beta)$ and independently $Y \sim \Gamma(\alpha_y, \beta)$. Let $Z = X + Y$. Then $Z \sim \Gamma((\alpha_x + \alpha_y), \beta)$.
Proof:

$$\begin{aligned} M_Z(t) &= \left(\frac{1}{1-\beta t}\right)^{\alpha_x} \left(\frac{1}{1-\beta t}\right)^{\alpha_y} \\ &= \left(\frac{1}{1-\beta t}\right)^{\alpha_x+\alpha_y} \end{aligned}$$

Remember that

- If $\alpha = 1$ we have an exponential with parameter β.
- If $\alpha = n/2$ and $\beta = 2$, we have a $\chi^2(n)$ (with n d.f.). The above result states that $\chi^2(n_1) + \chi^2(n_2) = \chi^2(n_1 + n_2)$.