Transformations of Two Random Variables

Problem : (X, Y) is a bivariate rv. Find the distribution of $Z = g(X, Y)$.

- The very 1st step: specify the support of Z.
- X, Y are discrete – straightforward; see Example 0(a)(b) from `Transformation_of_Several_Random_Variables.pdf`.
- X, Y are continuous
  - The CDF approach (the basic, off-the-shelf method)
  - Special formula (convolution) for $Z = X + Y$
  - MGF approach for sums of multiple independent rvs.

Examples of the CDF Approach

Example 3(f) from `Note_Mixture_Joint_0830.pdf`

Example (p83, Exercise 2.1.6) Let $f(x, y) = e^{-x-y}$, $0 < x, y < \infty$, zero elsewhere, be the pdf of X, Y.

a) Find the pdf of $Z = X + Y$.

$$\begin{aligned} F(Z \le z) &= \mathbb{P}(X + Y \le z) = \int_0^z dx \int_0^{z-x} e^{-x-y} dy \\ &= 1 - e^{-z} - ze^{-z}, \quad z > 0 \\ f_Z(z) &= ze^{-z}, \quad z > 0, \quad \text{i.e., } Z \sim \text{Ga}(2, 1) \end{aligned}$$

b) Find the pdf of $W = 2X + Y$.

$$\begin{aligned} F(W \le w) &= \mathbb{P}(2X + Y \le w) = \int_0^{w/2} dx \int_0^{w-2x} e^{-x-y} dy \\ &= 1 + e^{-w} - 2e^{-w/2}, \quad w > 0 \\ f_W(w) &= e^{-w/2} - e^{-w}, \quad w > 0. \end{aligned}$$

c) Find the pdf of $V = Y/X$.

$$\begin{aligned} F(V \le v) &= \mathbb{P}(Y \le vX) = \int_0^{\infty} dx \int_0^{vx} e^{-x-y} dy \\ &= 1 - \frac{1}{v+1}, \quad v > 0 \\ f_V(v) &= \frac{1}{(v+1)^2}, \quad v > 0 \end{aligned}$$

d) Find the pdf of $U = Y/(X+Y)$.

$$\begin{aligned} F(U \le u) &= \mathbb{P}(Y \le \frac{u}{1-u} X) = \int_0^\infty dx \int_0^{ux/(1-u)} e^{-x-y} dy \\ &= u, \quad 0 < u < 1 \\ f_U(u) &= 1, \quad 0 < u < 1. \end{aligned}$$

e) Find the pdf of $T = X - Y$.

$$\begin{aligned} F(T \le t) &= \mathbb{P}(X \le t + Y) = \int_0^\infty dy \int_0^{t+y} e^{-x-y} dx = 1 - e^{-t}/2, \ t > 0 \\ f_T(t) &= e^{-t}/2 \quad t > 0; \\ F(T \le t) &= \mathbb{P}(X \le t + Y) = \int_{-t}^\infty dy \int_0^{t+y} e^{-x-y} dx = e^{t}/2, \ t < 0 \\ f_T(t) &= e^{t}/2 \quad t < 0. \end{aligned}$$

That is, $f_T(t) = e^{-|t|}/2$ (Double Exponential).

The Convolution Formula

- Suppose X, Y are discrete, the pmf for $W = X + Y$ is given by

$$p_W(w) = \sum_{(x,y):x+y=w} p(x, y) = \sum_x p(x, w - y). \qquad (1)$$

- What if X, Y are continuous with pdf $f(x, y)$? Here is the guess: follow eq (1) but replace sum with integral

$$f_W(w) = \int_{-\infty}^{\infty} f(x, w - x)dx.$$

Our guess turns out to be correct; the rigorous proof is given on the next slide.

Let X, Y be continuous random variables with joint pdf $f(x, y)$. Then the pdf for $W = X + Y$ is given by

$$f_W(w) = \int_{-\infty}^{\infty} f(x, w-x)dx = \int_{-\infty}^{\infty} f(w-y, y)dy.$$

Proof:

$$\begin{aligned} F_W(w) &= \mathbb{P}(X+Y \le w) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{w-x} f(x,y)dy \\ &\quad \text{Change-of-variables: } y = u - x, \quad \text{i.e., } u = x + y. \\ &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{w} f(x, u-x)du \\ &= \int_{-\infty}^{w} du \int_{-\infty}^{\infty} f(x, u-x)dx \\ f_W(w) &= \frac{dF_W(w)}{dw} = \int_{-\infty}^{\infty} f(x, w-x)dx \end{aligned}$$

Don't forget about the range of w and the range of $(w-x)$!

- Example 2.1.6 (a)(Revisit): $X, Y \sim$ Exp(1) and they are independent. Find the pdf of $Z = X + Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, z{-}x)dx = \int_0^z e^{-x}e^{-(z-x)}dx = ze^{-z}, \quad Z \sim \text{Ga}(2,1).$$

Note that the joint pdf $f(x,y) = f_X(x)f_Y(y)$ is non-zero when $x, y > 0$, that is, $f(x, z-x) > 0$ only if $0 < x < z$.

- Example 2 from `Convolution2.pdf`.

The MGF Approach

If X, Y are independent, then[a]

$$\begin{aligned} M_{X+Y}(t) &= \mathbb{E}e^{t(X+Y)} = \mathbb{E}e^{tX} \times e^{tY} \\ &= (\mathbb{E}e^{tX})(\mathbb{E}e^{tY}) = M_X(t)M_Y(t). \end{aligned}$$

- Example 3 from `Convolution2.pdf`.
- Example 2.1.6 (a)(Revisit): $X, Y \sim$ Exp(1) and they are independent. Find the pdf of $Z = X + Y$.

[a] Don't confuse the equality above with $M_{XY}(t_1, t_2) = M_X(t_1)M_Y(t_2)$ for independent X, Y.

Additivity of Random Variables

If X, Y are independent,

- $X \sim \mathsf{Bin}(n_1, p), Y \sim \mathsf{Bin}(n_2, p) \Longrightarrow X + Y \sim \mathsf{Bin}(n_1 + n_2, p)$
- $X \sim \mathsf{Po}(\lambda_1), Y \sim \mathsf{Po}(\lambda_2) \Longrightarrow X + Y \sim \mathsf{Po}(\lambda_1 + \lambda_2)$
- $X \sim \mathsf{NB}(r_1, p), Y \sim \mathsf{NB}(r_2, p) \Longrightarrow X + Y \sim \mathsf{NB}(r_1 + r_2, p)$[a]
- $X \sim \mathsf{Geo}(p), Y \sim \mathsf{Geo}(p) \Longrightarrow X + Y \sim \mathsf{NB}(r = 2, p)$
- $X \sim \mathsf{Ga}(\alpha_1, \beta), Y \sim \mathsf{Ga}(\alpha_2, \beta) \Longrightarrow X + Y \sim \mathsf{Ga}(\alpha_1 + \alpha_2, \beta)$
- $X \sim \mathsf{Ex}(\lambda), Y \sim \mathsf{Ex}(\lambda) \Longrightarrow X + Y \sim \mathsf{Ga}(2, 1/\lambda)$
- $X \sim \chi^2(r_1), Y \sim \chi^2(r_2) \Longrightarrow X + Y \sim \chi^2(r_1 + r_2)$
- $X \sim \mathsf{No}(\mu_1, \sigma_1^2), Y \sim \mathsf{No}(\mu_2, \sigma_2^2)$

  $$\Longrightarrow X + Y \sim \mathsf{No}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

---
[a] NB = Negative Binomial.