Chapter 4
Multiple Random Variables
Generating Functions and Conditional Expectation

Outline

Probability Generating Functions

Moment Generating Functions

Conditional Mass Functions

Conditional Density Functions

Conditional Expectation

- Discrete Random Variables
- Continuous Random Variables

Conditional Variance

Probability Generating Functions

If X_1 and X_2 are independent discrete random variables, then the probability generating function for $X_1 + X_2$,

$$\rho_{X_1+X_2}(z) = E[z^{X_1+X_2}] = E[z^{X_1} z^{X_2}] = Ez^{X_1} \cdot Ez^{X_2} = \rho_{X_1}(z) \cdot \rho_{X_2}(z).$$

For a fair die, X

$$\rho_X(z) = \frac{1}{6}(z + z^2 + z^3 + z^4 + z^5 + z^6)$$

Thus, the generating function for the sum $S_n = X_1 + \cdots + X_n$ on n die is $\rho_X(z)^n$ and $f_{S_n}(x) = P\{S_n = x\}$ is the coefficient of z^x in the expansion of $\rho_X(z)^n$.

The generating function and its use in this case was introduced by Leonhard Euler.

Probability Generating Functions

Example. Let X_1 and X_2 be independent binomial random variables with common success parameter p. Assume n_1 trials for X_1 and n_2 trials for X_2. Then

$$\rho_{X_1+X_2}(z) = ((1-p)+pz)^{n_1} \cdot ((1-p)+pz)^{n_2} = ((1-p)+pz)^{n_1+n_2}.$$

Consequently, $X_1 + X_2$ is a binomial random variable with parameters p and $n_1 + n_2$.

Example. Let X_1 and X_2 be independent Poisson random variable with respective parameter values λ_1 and λ_2. Then,

$$\rho_{X_1+X_2}(z) = \exp(\lambda_1(z-1)) \cdot \exp(\lambda_2(z-1)) = \exp((\lambda_1+\lambda_2)(z-1)).$$

Consequently, $X_1 + X_2$ is a Poisson random variable with parameters $\lambda_1 + \lambda_2$.

Moment Generating Functions

A similar identity holds for the moment generating function for the sum of independent continuous random variables X_1 and X_2.

$$M_{X_1+X_2}(t) = E[e^{t(X_1+X_2)}] = E[e^{tX_1}e^{tX_2}] = Ee^{tX_1} \cdot Ee^{tX_2} = M_{X_1}(t) \cdot M_{X_2}(t).$$

The cumulant generating function,

$$K_{X_1+X_2}(t) = K_{X_1}(t) + K_{X_2}(t).$$

Also,

$$M_{\alpha+\beta X}(t) = E[e^{t(\alpha+\beta X)}] = e^{t\alpha} M_X(\beta t)$$

and

$$K_{\alpha+\beta X}(t) = t\alpha + K_X(\beta t)$$

Moment Generating Functions

Example. Let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ be independent normal random variables. Recall that the cumulant generating function for a standard normal, Z,

$$K_Z(t) = \frac{1}{2}t^2.$$

Thus, $X_i = \mu_i + \sigma_i Z_i$, for independent standard normals Z_1 and Z_2.

$$K_{X_i}(t) = \mu_i t + \frac{\sigma_i^2}{2}t^2$$

Thus,

$$K_{X_1+X_2}(t) = (\mu_1 + \mu_2)t + \frac{\sigma_1^2 + \sigma_1^2}{2}t^2$$

and $X_1 + X_2$ is a normal random variable with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.

Moment Generating Functions

Example. Let $X_1 \sim \Gamma(\alpha_1, \beta)$ and $X_2 \sim \Gamma(\alpha_2, \beta)$ be independent gamma random variables. For $t < \beta$, the moment generating function,

$$
\begin{aligned}
M_{X_i}(t) &= \frac{\beta^{\alpha_i}}{\Gamma(\alpha_i)} \int_0^\infty e^{tx} x^{\alpha_i} e^{-\beta x}\, dx = \frac{\beta^{\alpha_i}}{\Gamma(\alpha_i)} \int_0^\infty x^{\alpha_i} e^{-(\beta - t)x}\, dx. \\
&= \frac{\beta^{\alpha_i}}{(\beta - t)^{\alpha_i}} \cdot \frac{(\beta - t)^{\alpha_i}}{\Gamma(\alpha_i)} \int_0^\infty x^{\alpha_i} e^{-(\beta - t)x}\, dx. \\
&= \left(\frac{\beta}{\beta - t}\right)^{\alpha_i} \cdot 1 = \left(1 - \frac{t}{\beta}\right)^{-\alpha_i}
\end{aligned}
$$

and $M_{X_1+X_2}(t) = \left(1 - \frac{t}{\beta}\right)^{-\alpha_1} \left(1 - \frac{t}{\beta}\right)^{-\alpha_2} = \left(1 - \frac{t}{\beta}\right)^{-(\alpha_1+\alpha_2)}$, showing that $X_1 + X_2 \sim \Gamma(\alpha_1 + \alpha_2, \beta)$

Conditional Mass Functions

If X_1 and X_2 are discrete random variables, then we define the conditional mass function

$$f_{X_2|X_1}(x_2|x_1) = P\{X_2 = x_2|X_1 = x_1\} = \frac{P\{X_2 = x_2, X_1 = x_1\}}{P\{X_1 = x_1\}} = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)}.$$

Exercise. Check that $\sum_{x_2} f_{X_2|X_1}(x_2|x_1) = 1$.

$$\sum_{x_2} f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1).$$

Now divide by $f_{X_1}(x_1)$.

Exercise. If X and X_2 are independent, then $f_{X_2|X_1}(x_2|x_1) = f_{X_2}(x_2)$.

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)} = \frac{f_{X_1}(x_1)f_{X_2}(x_2)}{f_{X_1}(x_1)} = f_{X_2}(x_2)$$

Conditional Mass Functions

For X_1 and X_2 having joint mass function and conditional mass function

		x_1				
		0	1	2	3	4
x_2	0	0.02	0.02	0	0.10	0
	1	0.02	0.04	0.10	0	0
	2	0.02	0.06	0	0.10	0
	3	0.02	0.08	0.10	0	0.05
	4	0.02	0.10	0	0.10	0.05
	$f_{X_1}(x_1)$	0.10	0.30	0.20	0.30	0.10

		x_1				
		0	1	2	3	4
x_2	0	0.20	0.067	0	0.333	0
	1	0.20	0.133	0.50	0	0
	2	0.20	0.200	0	0.333	0
	3	0.20	0.267	0.50	0	0.50
	4	0.20	0.333	0	0.333	0.50

Conditional Density Functions

We define a conditional cumulative distribution function as a limit

$$F_{X_2|X_1}(x_2|x_1) = \lim_{\Delta x_1 \to 0} P\{X_2 \le x_2 | x_1 < X_1 \le x_1 + \Delta x_1\}.$$

Note that

$$P\{X_2 \le x_2 | x_1 < X_1 \le x_1 + \Delta x_1\} = \frac{P\{X_2 \le x_2, x_1 < X_1 \le x_1 + \Delta x_1\}}{P\{x_1 < X_1 \le x_1 + \Delta x_1\}}$$

$$= \frac{F_{X_1,X_2}(x_1 + \Delta x_1, x_2) - F_{X_1,X_2}(x_1, x_2)}{F_{X_1}(x_1 + \Delta x_1) - F_{X_1}(x_1)} \approx \frac{(\partial F_{X_1,X_2}(x_1, x_2)/\partial x_1)\Delta x_1}{f_{X_1}(x_1)\Delta x_1}$$

Thus,

$$F_{X_2|X_1}(x_2|x_1) = \frac{1}{f_{X_1}(x_1)} \frac{\partial}{\partial x_1} F_{X_1,X_2}(x_1, x_2).$$

The conditional density function

$$f_{X_2|X_1}(x_2|x_1) = \frac{1}{f_{X_1}(x_1)} \frac{\partial}{\partial x_2} F_{X_2|X_1}(x_2|x_1) = \frac{1}{f_{X_1}(x_1)} \frac{\partial^2}{\partial x_1 \partial x_2} F_{X_1,X_2}(x_1, x_2) = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)}$$

Conditional Density Functions

Exercise. Check that $\int_{-\infty}^{\infty} f_{X_2|X_1}(x_2|x_1)\, dx_2 = 1$.

If X and X_2 are independent, then $f_{X_2|X_1}(x_2|x_1) = f_{X_2}(x_2)$.

Let (X_1, X_2) have joint density

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \frac{4}{5}(x_1x_2 + x_1 + x_2) & \text{for } 0 \le x_1 \le 1, 0 \le x_2 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

The marginal density for X_1 is

$$f_{X_1}(x_1) = \frac{4}{5}\left(\frac{3}{2}x_1 + \frac{1}{2}\right)$$

and the conditional density

$$f_{X_2|X_1}(x_2|x_1) = \begin{cases} \frac{2(x_1x_2+x_1+x_2)}{3x_1+1} & \text{for } 0 \le x_2 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Bivariate Normal Density

For $-1 < \rho < 1$, let (Z_1, Z_2) have joint density

$$f_{Z_1,Z_2}(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp -\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2).$$

Then, the marginal density for Z_1

$$\begin{aligned}
f_{Z_1}(z_1) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp -\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\, dz_2 \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp -\frac{1}{2(1-\rho^2)}((1-\rho^2)z_1^2 + (z_2 - \rho z_1)^2)\, dz_2 \\
&= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} \exp -\frac{1}{2(1-\rho^2)}(z_2 - \rho z_1)^2\, dz_2 \\
&= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp -\frac{1}{2}u^2\, du, \qquad u = \frac{z_2 - \rho z_1}{\sqrt{1-\rho^2}} \\
&= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2}.
\end{aligned}$$

Bivariate Normal Density

Thus, Z_1 is a standard normal random variable. By a symmetric calculation, Z_2 is also a standard normal.

$$f_{Z_1,Z_2}(z_1,z_2) = \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp -\frac{1}{2(1-\rho^2)}(z_2-\rho z_1)^2.$$

The conditional density

$$f_{Z_2|Z_1}(z_2|z_1) = \frac{f_{Z_1,Z_2}(z_1,z_2)}{f_{Z_1}(z_1)} = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp -\frac{1}{2(1-\rho^2)}(z_2-\rho z_1)^2$$

and Z_2 conditioned on Z_1 taking the value z_1 is normal, mean ρz_1 and variance $1-\rho^2$.

Conditional Expectation

We will take an abstract approach to the definition to conditional expectation and see that it leads to the anticipated formula.

We would like the conditional expectation, $E[g(X_1, X_2)|X_1]$ to be a function of X_1 and to have the same averages whenever we average over an event that can be defined by X_1.

Definition. Call

$$h(X_1) = E[g(X_1, X_2)|X_1]$$

provided that for every measurable set B,

$$E[h(X_1)I_B(X_1)] = E[g(X_1, X_2)I_B(X_1)]$$

Discrete Random Variables

Take $B = \{x_1\}$, then

$$\begin{aligned}
E[h(X_1)I_{\{x_1\}}(X_1)] &= E[g(X_1, X_2)I_{\{x_1\}}(X_1)] \\
E[h(x_1)I_{\{x_1\}}(X_1)] &= \sum_{s_1, x_2} g(s_1, x_2)I_{\{x_1\}}(s_1)f_{X_1,X_2}(s_1, x_2) \\
h(x_1)f_{X_1}(x_1) &= \sum_{x_2} g(x_1, x_2)f_{X_1,X_2}(x_1, x_2) \\
h(x_1) &= \sum_{x_2} g(x_1, x_2)\frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)} = \sum_{x_2} g(x_1, x_2)f_{X_2|X_1}(x_2|x_1)
\end{aligned}$$

Thus, the conditional expectation $E[g(X_1, X_2)|X_1]$ is simply the expectation with respect to the conditional mass function. We shall often write

$$h(x_1) = E[g(X_1, X_2)|X_1 = x_1]$$

Discrete Random Variables

Example. We return to the joint mass function in the previous example. Here is the table for the conditional mass function $f_{X_2|X_1}(x_2|x_1)$ to find $E[X_2|X_1]$

		x_1				
		0	1	2	3	4
x_2	0	0.20	0.067	0	0.333	0
	1	0.20	0.133	0.50	0	0
	2	0.20	0.200	0	0.333	0
	3	0.20	0.267	0.50	0	0.50
	4	0.20	0.333	0	0.333	0.50
	$E[X_2\|X_1 = x_1]$	2.000	2.667	2.000	2.000	3.500

Continuous Random Variables

In this case we take $B = (-\infty, x_1]$, then

$$
\begin{aligned}
E[h(X_1) I_{(-\infty,x_1]}(X_1)] &= E[g(X_1, X_2) I_{(-\infty,x_1]}(X_1)] \\
\int_{-\infty}^{x_1} h(s_1) f_{X_1}(s_1)\ ds_1 &= \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} g(s_1, x_2) f_{X_1,X_2}(s_1, x_2)\ dx_2 ds_1 \\
h(x_1) f_{X_1}(x_1) &= \int_{-\infty}^{\infty} g(x_1, x_1) f_{X_1,X_2}(x_1, x_2)\ dx_2 \\
h(x_1) &= \int_{-\infty}^{\infty} g(x_1, x_2) \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)}\ dx_2 = \int_{-\infty}^{\infty} g(x_1, x_2) f_{X_2|X_1}(x_2|x_1)\ dx_2
\end{aligned}
$$

As before the conditional expectation $E[g(X_1, X_2)|X_1]$ is simply the expectation with respect to the conditional density function. Again, we write

$$h(x_1) = E[g(X_1, X_2)|X_1 = x_1]$$

Continuous Random Variables

For the example above

$$\begin{aligned}
E[X_2|X_1 = x_1] &= \int_{-\infty}^{\infty} x_2 f_{X_2|X_1}(x_2|x_1)\,dx_2 = \int_0^1 x_2 \frac{2(x_1x_2 + x_1 + x_2)}{3x_1+1}\,dx_2 \\
&= \frac{2}{3x_1+1}\left(\frac{1}{3}x_1x_2^3 + \frac{1}{2}x_1x_2^2 + \frac{1}{3}x_2^3\right)\Big|_0^1 = \frac{5x_1+2}{3(3x_1+1)}. \\
E[X_2|X_1] &= \frac{5X_1+2}{3(3X_1+1)}.
\end{aligned}$$

and

$$\begin{aligned}
E[X_2^2|X_1 = x_1] &= \int_{-\infty}^{\infty} x_2^2 f_{X_2|X_1}(x_2|x_1)\,dx_2 = \int_0^1 x_2^2 \frac{2(x_1x_2 + x_1 + x_2)}{3x_1+1}\,dx_2 \\
&= \frac{2}{3x_1+1}\left(\frac{1}{4}x_1x_2^4 + \frac{1}{3}x_1x_2^3 + \frac{1}{4}x_2^4\right)\Big|_0^1 = \frac{7x_1+3}{6(3x_1+1)}.
\end{aligned}$$

Conditional Variance

Conditional variance is the variance based on the conditional distribution given that $X_1 = x_1$

$$\mathrm{Var}(X_2|X_1) = E[(X_2 - E[X_2|X_1])^2|X_1] = E[X_2^2|X_1] - (E[X_2|X_1])^2.$$

Example For the discrete distribution example,

$$\mathrm{Var}(X_2|X_1 = 2) = E[X_2^2|X_1 = 2] - (E[X_2|X_1 = 2])^2 = 5 - 4 = 1.$$

For the continuous distribution example,

$$\begin{aligned}\mathrm{Var}(X_2|X_1) &= E[X_2^2|X_1] - (E[X_2|X_1])^2 = \frac{7X_1+3}{6(3X_1+1)} - \left(\frac{5X_1+2}{3(3X_1+1)}\right)^2 \\ &= \frac{3(7X_1+3)(3X_1+1) - 2(5X_1+2)^2}{18(3X_1+1)^2} \\ &= \frac{13X_1^2 + 8X_1 + 1}{18(3X_1+1)^2}\end{aligned}$$

Conditional Density Functions

Exercise. Let (X_1, X_2) have joint density $f_{(X_1,X_2)}(x_1, x_2) = e^{-x_2},\ 0 < x_1 < x_2 < \infty$. Then the marginal density

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\ dx_2 = \int_{x_1}^{\infty} e^{-x_2}\ dx_2 = e^{-x_1}.$$

Thus, X_1 is an $Exp(1)$ random variable. The conditional density is

$$f_{X_2|X_1}(x_2|x_1) = \begin{cases} e^{-(x_2-x_1)}, & \text{if } x_1 < x_2, \\ 0 \quad , & \text{if } x_2 \geq x_1. \end{cases}$$

Thus, given that $X_1 = x_1$, X_2 is equal to x_1 plus an $Exp(1)$ random variable. Thus, $E[X_2|X_1] = X_1 + 1$ and $\text{Var}(X_2|X_1) = 1$. Consequently, $EX_2 = 2$.