Chapter 4
Multivariate distributions

$k \geq 2$

Multivariate Distributions

All the results derived for the bivariate case can be generalized to *n* RV.

The joint CDF of $X_1, X_2, \ldots, X_k$ will have the form:

$P(x_1, x_2, \ldots, x_k)$ when the RVs are discrete

$F(x_1, x_2, \ldots, x_k)$ when the RVs are continuous

Joint Probability Function

Definition: Joint Probability Function

Let $X_1, X_2, \ldots, X_k$ denote k discrete random variables, then

$$p(x_1, x_2, \ldots, x_k)$$

is joint probability function of $X_1, X_2, \ldots, X_k$ if

1. $0 \leq p(X_1, \ldots, X_n) \leq 1$
2. $\sum_{x \in A}^{n} \ldots \sum_{x \in A}^{n} p(X_1, \ldots, X_n) = 1$
3. $P[(X_1, \ldots, X_n) \in A] = \sum_{x \in A}^{n} \ldots \sum_{x \in A}^{n} p(X_1, \ldots, X_n)$

Joint Density Function

Definition: Joint density function

Let $X_1, X_2, \ldots, X_k$ denote k continuous random variables, then

$$f(x_1, x_2, \ldots, x_k) = \delta^n/\delta x_1, \delta x_2, \ldots, \delta x_k \, F(x_1, x_2, \ldots, x_k)$$

is the joint density function of $X_1, X_2, \ldots, X_k$ if

1. $0 \leq f(X_1, \ldots, X_n)$
2. $\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1$
3. $P[(x_1, \ldots, x_n) \in A = \int_{A}^{\infty} \ldots \int_{A}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1$

Example: *The Multinomial distribution*

Suppose that we observe an experiment that has k possible outcomes $\{O_1, O_2, \ldots, O_k\}$ independently n times.

Let $p_1, p_2, \ldots, p_k$ denote probabilities of $O_1, O_2, \ldots, O_k$ respectively.

Let X_i denote the number of times that outcome O_i occurs in the n repetitions of the experiment.

Then the joint probability function of the random variables $X_1, X_2, \ldots, X_k$ is

$$p(X_1, \ldots, X_n) = \frac{n!}{x_1! \ldots x_n!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

Example: *The Multinomial distribution*

Note: $p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$

is the probability of a sequence of length n containing

x_1 outcomes O_1

x_2 outcomes O_2

…

x_k outcomes O_k

$$\binom{n}{x_1 \ldots x_k} = \frac{n!}{x_1! \ldots x_k!}$$

is the number of ways of choosing the positions for the x_1 outcomes O_1, x_2 outcomes O_2, …, x_k outcomes O_k

Example: *The Multinomial distribution*

$$\binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-x_1-x_2}{x_3}\cdots\binom{x_k}{x_k}$$

$$=\left(\frac{n!}{x_1!(n-x_1)!}\right)\left(\frac{(n-x_1)!}{x_2!(n-x_1-x_2)!}\right)\left(\frac{(n-x_1-x_2)!}{x_3!(n-x_1-x_2-x_3)!}\right)\cdots$$

$$=\frac{n!}{x_1!\ldots x_k!}$$

$$p(X_1,\ldots,X_n)=\frac{n!}{x_1!\ldots x_n!}p_1^{x_1}p_2^{x_2}\ldots p_k^{x_k}$$

$$=\binom{n}{x_1\ldots x_k}p_1^{x_1}p_2^{x_2}\ldots p_k^{x_k}$$

This is called the **Multinomial** distribution

Example: *The Multinomial distribution*

Suppose that an earnings announcements has three possible outcomes:

O_1 – Positive stock price reaction – (30% chance)
O_2 – No stock price reaction – (50% chance)
O_3 - Negative stock price reaction – (20% chance)

Hence $p_1 = 0.30$, $p_2 = 0.50$, $p_3 = 0.20$.

Suppose today 4 firms released earnings announcements ($n = 4$). Let X = the number that result in a positive stock price reaction, Y = the number that result in no reaction, and Z = the number that result in a negative reaction.

Find the distribution of X, Y and Z. Compute $P[X + Y \geq Z]$

$$p(x,y,z)=\frac{4!}{x!y!z!}(0.30)^x(0.50)^y(0.20)^z \quad x+y+z=4$$

Table: $p(x,y,z)$

x	y	z				
		0	1	2	3	4
0	0	0	0	0	0	0.0016
0	1	0	0	0	0.0160	0
0	2	0	0	0.0600	0	0
0	3	0	0.1000	0	0	0
0	4	0.0625	0	0	0	0
1	0	0	0	0	0.0096	0
1	1	0	0	0.0720	0	0
1	2	0	0.1800	0	0	0
1	3	0.1500	0	0	0	0
1	4	0	0	0	0	0
2	0	0	0	0.0216	0	0
2	1	0	0.1080	0	0	0
2	2	0.1350	0	0	0	0
2	3	0	0	0	0	0
2	4	0	0	0	0	0
3	0	0	0.0216	0	0	0
3	1	0.0540	0	0	0	0
3	2	0	0	0	0	0
3	3	0	0	0	0	0
3	4	0	0	0	0	0
4	0	0.0081	0	0	0	0
4	1	0	0	0	0	0
4	2	0	0	0	0	0
4	3	0	0	0	0	0
4	4	0	0	0	0	0

$P[X + Y \geq Z]$

$= 0.9728$

x	y	z				
		0	1	2	3	4
0	0	0	0	0	0	0.0016
0	1	0	0	0	0.0160	0
0	2	0	0	0.0600	0	0
0	3	0	0.1000	0	0	0
0	4	0.0625	0	0	0	0
1	0	0	0	0	0.0096	0
1	1	0	0	0.0720	0	0
1	2	0	0.1800	0	0	0
1	3	0.1500	0	0	0	0
1	4	0	0	0	0	0
2	0	0	0	0.0216	0	0
2	1	0	0.1080	0	0	0
2	2	0.1350	0	0	0	0
2	3	0	0	0	0	0
2	4	0	0	0	0	0
3	0	0	0.0216	0	0	0
3	1	0.0540	0	0	0	0
3	2	0	0	0	0	0
3	3	0	0	0	0	0
3	4	0	0	0	0	0
4	0	0.0081	0	0	0	0
4	1	0	0	0	0	0
4	2	0	0	0	0	0
4	3	0	0	0	0	0
4	4	0	0	0	0	0

Example: *The Multivariate Normal distribution*

Recall the univariate normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

the bivariate normal distribution

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{x-\mu_y}{\sigma_y}\right) + \left(\frac{x-\mu_y}{\sigma_y}\right)^2\right]}$$

Example: *The Multivariate Normal distribution*

Note: We can have a more compact joint using linear algebra:

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)}\left((\frac{x_1-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2}) + (\frac{x_2-\mu_2}{\sigma_2})^2\right)^2\right]$$

$$= \frac{1}{2\pi(|\Sigma|)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

(1) Determine the inverse and determinant of Σ (the covariance matrix)

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \Rightarrow |\Sigma| = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2(1-\frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}) = \sigma_1^2\sigma_2^2(1-\rho^2)$$

$$\Rightarrow \Sigma^{-1} = \frac{1}{|\Sigma|}\begin{bmatrix} \sigma_2^2 & -\sigma_{21} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

The bivariate normal distribution

(2) Write a quadratic form for $Q(x_1,x_2)$:

$$Q(x_1,x_2)=\frac{\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}}{1-\rho^2}$$

$$Q(x_1,x_2)=(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

$$=\left[(x_1-\mu_1)\quad(x_2-\mu_2)\right]\frac{1}{|\Sigma|}\begin{bmatrix}\sigma_2^2 & -\sigma_{21}\\ -\sigma_{12} & \sigma_1^2\end{bmatrix}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}=$$

$$=\frac{1}{|\Sigma|}\left[(x_1-\mu_1)\sigma_2^2-(x_2-\mu_2)\sigma_{12}\quad(x_1-\mu_1)\sigma_{21}-(x_2-\mu_2)\sigma_1^2\right]\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}$$

$$=\frac{1}{|\Sigma|}((x_1-\mu_1)^2\sigma_2^2-2\sigma_{21}(x_1-\mu_1)(x_2-\mu_2)+(x_2-\mu_2)^2\sigma_1^2)$$

$$=\frac{((x_1-\mu_1)^2\sigma_2^2-2\sigma_{21}(x_1-\mu_1)(x_2-\mu_2)+(x_2-\mu_2)^2\sigma_1^2)}{\sigma_1^2\sigma_2^2(1-\rho^2)}$$

Example: *The Multivariate Normal distribution*

The *k-variate Normal distribution* is given by:

$$f(x_1,\ldots,x_k)=f(\mathbf{x})=\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}}e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

where

$$\mathbf{x}=\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_k\end{bmatrix}\qquad \boldsymbol{\mu}=\begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_k\end{bmatrix}\qquad \Sigma=\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k}\\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2k}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_{kk}\end{bmatrix}$$

Marginal joint probability function

Definition: Marginal joint probability function

Let $X_1, X_2, \ldots, X_q, X_{q+1} \ldots, X_k$ denote k discrete random variables with joint probability function

$$p(x_1, x_2, \ldots, x_q, x_{q+1} \ldots, x_k)$$

then the *marginal joint probability function* of $X_1, X_2, \ldots, X_q$ is

$$p_{12\ldots q}\left(x_1,\ldots,x_q\right) = \sum_{x_{q+1}}\cdots\sum_{x_n} p\left(x_1,\ldots,x_n\right)$$

When $X_1, X_2, \ldots, X_q, X_{q+1} \ldots, X_k$ is continuous, then the *marginal joint density function* of $X_1, X_2, \ldots, X_q$ is

$$f_{12\ldots q}\left(x_1,\ldots,x_q\right) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f\left(x_1,\ldots,x_n\right)dx_{q+1}\ldots dx_n$$

Conditional joint probability function

Definition: Conditional joint probability function

Let $X_1, X_2, \ldots, X_q, X_{q+1} \ldots, X_k$ denote k discrete random variables with joint probability function

$$p(x_1, x_2, \ldots, x_q, x_{q+1} \ldots, x_k)$$

then the *conditional joint probability function* of $X_1, X_2, \ldots, X_q$ given $X_{q+1} = x_{q+1}, \ldots, X_k = x_k$ is

$$p_{1\ldots q|q+1\ldots k}\left(x_1,\ldots,x_q \middle| x_{q+1},\ldots,x_k\right) = \frac{p\left(x_1,\ldots,x_k\right)}{p_{q+1\ldots k}\left(x_{q+1},\ldots,x_k\right)}$$

For the continuous case, we have:

$$f_{1\ldots q|q+1\ldots k}\left(x_1,\ldots,x_q \middle| x_{q+1},\ldots,x_k\right) = \frac{f\left(x_1,\ldots,x_k\right)}{f_{q+1\ldots k}\left(x_{q+1},\ldots,x_k\right)}$$

Conditional joint probability function

Definition: Independence of sets of vectors

Let $X_1, X_2, \ldots, X_q, X_{q+1} \ldots, X_k$ denote k continuous random variables with joint probability density function

$f(x_1, x_2, \ldots, x_q, x_{q+1} \ldots, x_k)$

then the variables $X_1, X_2, \ldots, X_q$ are **independent** of $X_{q+1}, \ldots, X_k$ if

$$f\left(x_1,\ldots,x_k\right) = f_{1\ldots q}\left(x_1,\ldots,x_q\right) f_{q+1\ldots k}\left(x_{q+1},\ldots,x_k\right)$$

A similar definition for discrete random variables.

Conditional joint probability function

Definition: Mutual Independence

Let $X_1, X_2, \ldots, X_k$ denote k continuous random variables with joint probability density function

$f(x_1, x_2, \ldots, x_k)$

then the variables $X_1, X_2, \ldots, X_k$ are called **mutually independent** if

$$f\left(x_1,\ldots,x_k\right) = f_1\left(x_1\right) f_2\left(x_2\right) \ldots f_k\left(x_k\right)$$

A similar definition for discrete random variables.

Multivariate marginal pdfs - Example

Let X, Y, Z denote 3 jointly distributed random variable with joint density function then

$$f(x,y,z) = \begin{cases} K(x^2 + yz) & 0 \le x \le 1, 0 \le y \le 1, 0 \le z \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of K.

Determine the marginal distributions of X, Y and Z.

Determine the joint marginal distributions of

X, Y

X, Z

Y, Z

Multivariate marginal pdfs - Example

Solution: Determining the value of K.

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,dxdydz = \int_0^1\int_0^1\int_0^1 K(x^2 + yz)\,dxdydz$$

$$= K\int_0^1\int_0^1 \left[\frac{x^3}{3} + xyz\right]_{x=0}^{x=1} dydz = K\int_0^1\int_0^1 \left(\frac{1}{3} + yz\right) dydz$$

$$= K\int_0^1 \left[\frac{1}{3}y + z\frac{y^2}{2}\right]_{y=0}^{y=1} dz = K\int_0^1 \left(\frac{1}{3} + z\frac{1}{2}\right) dz$$

$$= K\left[\frac{z}{3} + \frac{z^2}{4}\right]_0^1 = K\left(\frac{1}{3} + \frac{1}{4}\right) = K\frac{7}{12} = 1 \qquad \text{if } K = \frac{12}{7}$$

Multivariate marginal pdfs - Example

The marginal distribution of X.

$$f_1(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,dy\,dz = \frac{12}{7}\int_0^1\int_0^1 \left(x^2 + yz\right) dy\,dz$$

$$= \frac{12}{7}\int_0^1 \left[x^2 y + \frac{y^2}{2}z\right]_{y=0}^{y=1} dz = \frac{12}{7}\int_0^1 \left(x^2 + \frac{1}{2}z\right) dz$$

$$= \frac{12}{7}\left[x^2 z + \frac{z^2}{4}\right]_0^1 = \frac{12}{7}\left(x^2 + \frac{1}{4}\right) \quad \text{for} \quad 0 \le x \le 1$$

Multivariate marginal pdfs - Example

The marginal distribution of X,Y.

$$f_{12}(x,y) = \int_{-\infty}^{\infty} f(x,y,z)\,dz = \frac{12}{7}\int_0^1 \left(x^2 + yz\right) dz$$

$$= \frac{12}{7}\left[x^2 z + y\frac{z^2}{2}\right]_{z=0}^{z=1}$$

$$= \frac{12}{7}\left(x^2 + \frac{1}{2}y\right) \quad \text{for} \quad 0 \le x \le 1, 0 \le y \le 1$$

Multivariate marginal pdfs - Example

Find the conditional distribution of:

1. Z given $X = x$, $Y = y$,
2. Y given $X = x$, $Z = z$,
3. X given $Y = y$, $Z = z$,
4. Y , Z given $X = x$,
5. X , Z given $Y = y$
6. X , Y given $Z = z$
7. Y given $X = x$,
8. X given $Y = y$
9. X given $Z = z$
10. Z given $X = x$,
11. Z given $Y = y$
12. Y given $Z = z$

Multivariate marginal pdfs - Example

The marginal distribution of X,Y.

$$f_{12}(x,y) = \frac{12}{7}\left(x^2 + \frac{1}{2}y\right) \quad \text{for} \quad 0 \le x \le 1, 0 \le y \le 1$$

Thus the conditional distribution of Z given $X = x, Y = y$ is

$$\frac{f(x,y,z)}{f_{12}(x,y)} = \frac{\frac{12}{7}(x^2 + yz)}{\frac{12}{7}\left(x^2 + \frac{1}{2}y\right)}$$

$$= \frac{x^2 + yz}{x^2 + \frac{1}{2}y} \quad \text{for} \quad 0 \le z \le 1$$

Multivariate marginal pdfs - Example

The marginal distribution of X.

$$f_1(x) = \frac{12}{7}\left(x^2 + \frac{1}{4}\right) \quad \text{for} \quad 0 \le x \le 1$$

Then, the conditional distribution of Y, Z given $X = x$ is

$$\frac{f(x, y, z)}{f_1(x)} = \frac{\frac{12}{7}\left(x^2 + yz\right)}{\frac{12}{7}\left(x^2 + \frac{1}{4}\right)}$$

$$= \frac{x^2 + yz}{x^2 + \frac{1}{4}} \quad \text{for} \quad 0 \le y \le 1, 0 \le z \le 1$$

Expectations for Multivariate Distributions

Definition: Expectation

Let $X_1, X_2, \ldots, X_n$ denote n jointly distributed random variable with joint density function

$$f(x_1, x_2, \ldots, x_n)$$

then

$$E\left[g\left(X_1, \ldots, X_n\right)\right]$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g\left(x_1, \ldots, x_n\right) f\left(x_1, \ldots, x_n\right) dx_1, \ldots, dx_n$$

Expectations for Multivariate Distributions - Example

Let X, Y, Z denote 3 jointly distributed random variable with joint density function then

$$f(x,y,z)=\begin{cases}\frac{12}{7}\left(x^2+yz\right) & 0\le x\le 1, 0\le y\le 1, 0\le z\le 1\\ 0 & \text{otherwise}\end{cases}$$

Determine E[XYZ].

Solution:

$$E[XYZ]=\int_0^1\int_0^1\int_0^1 xyz\frac{12}{7}\left(x^2+yz\right)dxdydz$$

$$=\frac{12}{7}\int_0^1\int_0^1\int_0^1\left(x^3yz+xy^2z^2\right)dxdydz$$

Expectations for Multivariate Distributions - Example

$$E[XYZ]=\int_0^1\int_0^1\int_0^1 xyz\frac{12}{7}\left(x^2+yz\right)dxdydz=\frac{12}{7}\int_0^1\int_0^1\int_0^1\left(x^3yz+xy^2z^2\right)dxdydz$$

$$=\frac{12}{7}\int_0^1\int_0^1\left[\frac{x^4}{4}yz+\frac{x^2}{2}y^2z^2\right]_{x=0}^{x=1}dydz=\frac{3}{7}\int_0^1\int_0^1\left(yz+2y^2z^2\right)dydz$$

$$=\frac{3}{7}\int_0^1\left[\frac{y^2}{2}z+2\frac{y^3}{3}z^2\right]_{y=0}^{y=1}dz=\frac{3}{7}\int_0^1\left(\frac{1}{2}z+\frac{2}{3}z^2\right)dz$$

$$=\frac{3}{7}\left[\frac{z^2}{4}+\frac{2z^3}{9}\right]_0^1=\frac{3}{7}\left(\frac{1}{4}+\frac{2}{9}\right)=\frac{3}{7}\left(\frac{17}{36}\right)=\frac{17}{84}$$

Some Rules for Expectations – Rule 1

1. $$E[X_i] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f(x_1, \dots, x_n)\, dx_1 \dots dx_n = \int_{-\infty}^{\infty} x_i f_i(x_i)\, dx_i$$

Thus you can calculate $E[X_i]$ either from the joint distribution of $X_1, \dots, X_n$ or the marginal distribution of X_i.

Proof:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f(x_1, \dots, x_n)\, dx_1, \dots, dx_n$$

$$= \int_{-\infty} x_i \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n)\, dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n \right] dx_i$$

$$= \int_{-\infty}^{\infty} x_i f_i(x_i)\, dx_i$$

Some Rules for Expectations – Rule 2

2. $$E[a_1 X_1 + \cdots + a_n X_n] = a_1 E[X_1] + \cdots + a_n E[X_n]$$

This property is called the *Linearity property*.

Proof:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (a_1 x_1 + \dots + a_n x_n) f(x_1, \dots, x_n)\, dx_1 \dots dx_n$$

$$= a_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 f(x_1, \dots, x_n)\, dx_1 \dots dx_n$$

$$+ a_n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_n f(x_1, \dots, x_n)\, dx_1 \dots dx_n$$

Some Rules for Expectations – Rule 3

3. **(The Multiplicative property)** Suppose $X_1, \dots, X_q$ are independent of $X_{q+1}, \dots, X_k$ then

$$E\left[g\left(X_1,\dots,X_q\right)h\left(X_{q+1},\dots,X_k\right)\right]$$
$$= E\left[g\left(X_1,\dots,X_q\right)\right]E\left[h\left(X_{q+1},\dots,X_k\right)\right]$$

In the simple case when $k = 2$, and g(X)=X & h(Y)=Y:

$$E\left[XY\right] = E\left[X\right]E\left[Y\right]$$

if X and Y are independent

Some Rules for Expectations – Rule 3

Proof: $E\left[g\left(X_1,\dots,X_q\right)h\left(X_{q+1},\dots,X_k\right)\right]$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(x_1,\dots,x_q\right)h\left(x_{q+1},\dots,x_k\right)f\left(x_1,\dots,x_k\right)dx_1\dots dx_n$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(x_1,\dots,x_q\right)h\left(x_{q+1},\dots,x_k\right)f_1\left(x_1,\dots,x_q\right)$$
$$f_2\left(x_{q+1},\dots,x_k\right)dx_1\dots dx_q dx_{q+1}\dots dx_k$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h\left(x_{q+1},\dots,x_k\right)f_2\left(x_{q+1},\dots,x_k\right)\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(x_1,\dots,x_q\right)\right.$$
$$\left. f_1\left(x_1,\dots,x_q\right)dx_1\dots dx_q\right]dx_{q+1}\dots dx_k$$

$$= E\left[g\left(X_1,\dots,X_q\right)\right]\times$$
$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h\left(x_{q+1},\dots,x_k\right)f_2\left(x_{q+1},\dots,x_k\right)dx_{q+1}\dots dx_k$$

Some Rules for Expectations – Rule 3

$$= E\left[g\left(X_1,\ldots,X_q\right)\right]\times$$

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h\left(x_{q+1},\ldots,x_k\right)f_2\left(x_{q+1},\ldots,x_k\right)dx_{q+1}\ldots dx_k$$

$$= E\left[g\left(X_1,\ldots,X_q\right)\right]E\left[h\left(X_{q+1},\ldots,X_k\right)\right]$$

Some Rules for Variance – Rule 1

1. $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)$

   where $\text{Cov}(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$

Proof:

$$\text{Var}(X+Y) = E\left[\left((X+Y)-\mu_{X+Y}\right)^2\right]$$

where $\mu_{X+Y} = E[X+Y] = \mu_X + \mu_Y$

Thus,

$$\text{Var}(X+Y) = E\left[\left((X+Y)-(\mu_X+\mu_Y)\right)^2\right]$$

$$= E\left[(X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2\right]$$

$$= \text{Var}(X) + 2\text{Cov}(X,Y) + \text{Var}(Y)$$

Some Rules for Variance – Rule 1

Note: If X and Y are independent, then

$$\text{Cov}(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$
$$= E[X-\mu_X]\,E[Y-\mu_Y]$$
$$= (E[X]-\mu_X)(E[Y]-\mu_Y) = 0$$

and $\quad \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

Some Rules for Variance – Rule 1 - ρ_{XY}

<u>Definition</u>**:** Correlation coefficient

For any two random variables X and Y then define the *correlation coefficient* ρ_{XY} to be:

$$\rho_{xy} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Thus $\text{Cov}(X,Y) = \rho_{XY}\sigma_X\sigma_Y$

and $\quad \text{Var}(X+Y) = \sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y$

$\qquad = \sigma_X^2 + \sigma_Y^2$ if X and Y are independent.

Some Rules for Variance – Rule 1 - ρ_{XY}

Recall $\rho_{xy} = \dfrac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \dfrac{Cov(X,Y)}{\sigma_X \sigma_Y}$

Property 1. If X and Y are independent, then $\rho_{XY} = 0$. (Cov(X,Y)=0.)

The converse is not necessarily true. That is, $\rho_{XY} = 0$ does not imply that X and Y are independent.

Example:

$y \backslash x$	6	8	10	$f_y(y)$
1	.2	0	.2	.4
2	0	.2	0	.2
3	.2	0	.2	.4
$f_x(x)$	.4	.2	.4	1

$E(X)=8$, $E(Y)=2$, $E(XY)=16$
$Cov(X,Y) = 16 - 8*2 = 0$

$P(X=6,Y=2)=0 \neq P(X=6)*P(Y=2)=.4*$
$*.2=.08 \Rightarrow X \& Y$ are not independent.

Some Rules for Variance – Rule 1 - ρ_{XY}

Property 2. $-1 \leq \rho_{XY} \leq 1$

and $|\rho_{XY}| = 1$ if there exists a and b such that

$$P[Y = bX + a] = 1$$

where $\rho_{XY} = +1$ if $b > 0$ and $\rho_{XY} = -1$ if $b < 0$

Proof: Let $U = X - \mu_X$ and $V = Y - \mu_Y$.

Let $g(b) = E\left[(V - bU)^2\right] \geq 0$ for all b.

We will pick b to minimize g(b).

$$g(b) = E\left[(V - bU)^2\right] = E\left[V^2 - 2bVU + b^2U^2\right]$$
$$= E\left[V^2\right] - 2bE\left[VU\right] + b^2E\left[U^2\right]$$

Some Rules for Variance – Rule 1 - ρ_{XY}

Taking first derivatives of g(*b*) w.r.t *b*

$$g(b) = E\left[(V - bU)^2\right] = E\left[V^2\right] - 2bE\left[VU\right] + b^2 E\left[U^2\right]$$

$$g'(b) = -2E\left[VU\right] + 2bE\left[U^2\right] = 0 \quad \Rightarrow \quad b = b_{\min} = \frac{E\left[VU\right]}{E\left[U^2\right]}$$

Since $g(b) \geq 0$, then $g(b_{\min}) \geq 0$

$$g(b_{\min}) = E\left[V^2\right] - 2b_{\min}E\left[VU\right] + b_{\min}^2 E\left[U^2\right]$$

$$= E\left[V^2\right] - 2\frac{E\left[VU\right]}{E\left[U^2\right]}E\left[VU\right] + \left(\frac{E\left[VU\right]}{E\left[U^2\right]}\right)^2 \mathrm{E}[U^2]$$

$$= E\left[V^2\right] - \frac{\left(E\left[VU\right]\right)^2}{E\left[U^2\right]} \geq 0$$

Some Rules for Variance – Rule 1 - ρ_{XY}

$$= E\left[V^2\right] - \frac{\left(E\left[VU\right]\right)^2}{E\left[U^2\right]} \geq 0$$

Thus, $$\frac{\left(E\left[VU\right]\right)^2}{E\left[U^2\right]E\left[V^2\right]} \leq 1$$

or $$\frac{\left(E\left[(X-\mu_X)(Y-\mu_Y)\right]\right)^2}{E\left[(X-\mu_X)^2\right]E\left[(Y-\mu_Y)^2\right]} = \rho_{XY}^2 \leq 1$$

$$\Rightarrow \quad -1 \leq \rho_{XY} \leq 1$$

$$\Rightarrow \quad \rho_{XY} = 1 \qquad \text{if } \sigma_{XY}^2 = \sigma_X^2\sigma_Y^2$$

Some Rules for Variance – Rule 1 - ρ_{XY}

Note: $g\left(b_{\min}\right) = E\left[V^2\right] - 2b_{\min} E\left[VU\right] + b_{\min}^2 E\left[U^2\right]$

$$= E\left[\left(V - b_{\min} U\right)^2\right] = 0$$

If and only if $\rho_{XY}^2 = 1$

This will be true if

$$P\left[\left(Y - \mu_Y\right) - b_{\min}\left(X - \mu_X\right) = 0\right] = 1$$

$$P\left[Y = b_{\min} X + a\right] = 1 \text{ where } a = \mu_Y - b_{\min}\mu_X$$

i.e., $$P\left[V - b_{\min} U = 0\right] = 1$$

Some Rules for Variance – Rule 1 - ρ_{XY}

- Summary:

$$-1 \leq \rho_{XY} \leq 1$$

and $\left|\rho_{XY}\right| = 1$ if there exists *a* and *b* such that

$$P\left[Y = bX + a\right] = 1$$

where $$b = b_{\min} = \frac{E\left[\left(X - \mu_X\right)\left(Y - \mu_X\right)\right]}{E\left[\left(X - \mu_X\right)^2\right]}$$

$$= \frac{\text{Cov}\left(X, Y\right)}{\text{Var}\left(X\right)} = \frac{\rho_{XY}\sigma_X\sigma_Y}{\sigma_X^2} = \rho_{XY}\frac{\sigma_Y}{\sigma_X}$$

and $$a = \mu_Y - b_{\min}\mu_X = \mu_Y - \rho_{XY}\frac{\sigma_Y}{\sigma_X}\mu_X$$

Some Rules for Variance – Rule 2

2. $\text{Var}(aX+bY)=a^2\text{Var}(X)+b^2\text{Var}(Y)+2ab\text{Cov}(X,Y)$

Proof

$$\text{Var}(aX+bY)=E\left[\left((aX+bY)-\mu_{aX+bY}\right)^2\right]$$

with $\mu_{aX+bY}=E[aX+bY]=a\mu_X+b\mu_Y$

Thus,

$$\begin{aligned}\text{Var}(aX+bY)&=E\left[\left((aX+bY)-(a\mu_X+b\mu_Y)\right)^2\right]\\&=E\left[a^2(X-\mu_X)^2+2ab(X-\mu_X)(Y-\mu_Y)+b^2(Y-\mu_Y)^2\right]\\&=a^2\text{Var}(X)+2ab\text{Cov}(X,Y)+b^2\text{Var}(Y)\end{aligned}$$

Some Rules for Variance – Rule 3

3. $\text{Var}(a_1X_1+\ldots+a_nX_n)=$

$$\begin{aligned}&a_1^2\text{Var}(X_1)+\ldots+a_n^2\text{Var}(X_n)+\\&\quad+2a_1a_2\text{Cov}(X_1,X_2)+\ldots+2a_1a_n\text{Cov}(X_1,X_n)\\&\quad\quad+2a_2a_3\text{Cov}(X_2,X_3)+\ldots+2a_2a_n\text{Cov}(X_2,X_n)\\&\quad\quad\quad+2a_{n-1}a_n\text{Cov}(X_{n-1},X_n)\end{aligned}$$

$$=\sum_{i=1}^n a_i^2\text{Var}(X_i)+2\sum_{i<j}\sum a_ia_j\text{Cov}(X_i,X_j)$$

$$=\sum_{i=1}^n a_i^2\text{Var}(X_i)\ \text{ if } X_1,\ldots,X_n \text{ are mutually independent}$$

The mean and variance of a Binomial RV

We have already computed this by other methods:

1. Using the probability function $p(x)$.
2. Using the moment generating function $m_X(t)$.

Now, we will apply the previous rules for mean and variances.

Suppose that we have observed n independent repetitions of a Bernoulli trial.

Let $X_1, \dots, X_n$ be n mutually independent random variables each having Bernoulli distribution with parameter p and defined by

$$X_i = \begin{cases} 1 & \text{if repetition } i \text{ is } \mathbf{S} \text{ (prob } = p) \\ 0 & \text{if repetition } i \text{ is } \mathbf{F} \text{ (prob } = q) \end{cases}$$

The mean and variance of a Binomial RV

$$\mu = E\left[X_i\right] = 1 \cdot p + 0 \cdot q = p$$

$$\sigma^2 = Var[X_i] = (1-p)^2 p + (0-p)^2 q = (1-p)^2 p + (0-p)^2 (1-p) =$$
$$= (1-p)\ (p - p^2 + p^2) = qp$$

• Now $X = X_1 + \dots + X_n$ has a Binomial distribution with parameters n and p. Then, X is the total number of successes in the n repetitions.

$$\mu_X = E\left[X_1\right] + \ldots + E\left[X_n\right] = p + \ldots + p = np$$

$$\sigma_X^2 = \text{var}\left[X_1\right] + \ldots + \text{var}\left[X_n\right] = pq + \ldots + pq = npq$$

Conditional Expectation

Definition: Conditional Joint Probability Function

Let $X_1, X_2, \ldots, X_q, X_{q+1} \ldots, X_k$ denote k continuous random variables with joint probability density function

$f(x_1, x_2, \ldots, x_q, x_{q+1} \ldots, x_k)$

then the **conditional** joint probability function of $X_1, X_2, \ldots, X_q$ given $X_{q+1} = x_{q+1}, \ldots, X_k = x_k$ is

$$f_{1\ldots q|q+1\ldots k}\left(x_1, \ldots, x_q \,\middle|\, x_{q+1}, \ldots, x_k\right) = \frac{f\left(x_1, \ldots, x_k\right)}{f_{q+1\ldots k}\left(x_{q+1}, \ldots, x_k\right)}$$

Definition: Conditional Joint Probability Function

Let $U = h(X_1, X_2, \ldots, X_q, X_{q+1} \ldots, X_k)$ then the *Conditional Expectation* of U given $X_{q+1} = x_{q+1}, \ldots, X_k = x_k$ is

$$E\left[U \middle| x_{q+1}, \ldots, x_k\right] =$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \ldots, x_k) f_{1\ldots q|q+1\ldots k}\left(x_1, \ldots, x_q \middle| x_{q+1}, \ldots, x_k\right) dx_1 \ldots dx_q$$

Note: This will be a function of $x_{q+1}, \ldots, x_k$.

• Let Y and X have a joint pdf f_{YX}. Then,

$$\mathrm{E}[y\,|\,x] = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x) dy \text{ is called } \textit{regression} \text{ of } y \text{ on } x.$$

• $\mathrm{E}[y\,|\,x]$ is a function of x. Very useful result.

Definition: Conditional Joint Probability Function

• Any random variable Y can be expressed as the conditional mean plus an error term, ε, defined as $\varepsilon = (y - \mathrm{E}[y\,|\,x])$:

$$\begin{aligned} y &= \mathrm{E}[y\,|\,x] + (y - \mathrm{E}[y\,|\,x]) \\ &= \mathrm{E}[y\,|\,x] + \varepsilon. \end{aligned}$$

Depending on $\mathrm{E}[y\,|\,x]$, we may have a linear model. The conditional mean is what researchers model. It is a function of x.

Example: In the CAPM, equilibrium expected excess returns (y) are only determined by expected excess market returns (x):

$$\mathrm{E}[r_{i,t} - r_f] = \beta_i\, \mathrm{E}[(r_{m,t} - r_f)].$$

Then,

$$r_{i,t} - r_f = \alpha_i + \beta_i\,(r_{m,t} - r_f) + \varepsilon_{i,t}, \qquad i = 1, \ldots, N \ \&\ t = 1, \ldots, T$$

Conditional Expectation of a Function - Example

Let X, Y, Z denote 3 jointly distributed RVs with joint density function then

$$f(x,y,z)=\begin{cases}\frac{12}{7}\left(x^2+yz\right) & 0\le x\le 1, 0\le y\le 1, 0\le z\le 1\\ 0 & \text{otherwise}\end{cases}$$

Determine the conditional expectation of $U = X^2 + Y + Z$ given $X = x$, $Y = y$.

Integration over z, gives us the marginal distribution of X,Y:

$$f_{12}(x,y)=\frac{12}{7}\left(x^2+\frac{1}{2}y\right) \quad \text{for} \quad 0\le x\le 1, 0\le y\le 1$$

Conditional Expectation of a Function - Example

Then, the conditional distribution of Z given $X = x$, $Y = y$ is:

$$\frac{f(x,y,z)}{f_{12}(x,y)}=\frac{\frac{12}{7}\left(x^2+yz\right)}{\frac{12}{7}\left(x^2+\frac{1}{2}y\right)}$$

$$=\frac{x^2+yz}{x^2+\frac{1}{2}y} \quad \text{for} \quad 0\le z\le 1$$

Conditional Expectation of a Function - Example

The conditional expectation of $U = X^2 + Y + Z$ given $X = x$, $Y = y$.

$$E\left[U \mid x, y\right] = \int_0^1 \left(x^2 + y + z\right) \frac{x^2 + yz}{x^2 + \frac{1}{2} y} dz$$

$$= \frac{1}{x^2 + \frac{1}{2} y} \int_0^1 \left(x^2 + y + z\right)\left(x^2 + yz\right) dz$$

$$= \frac{1}{x^2 + \frac{1}{2} y} \int_0^1 \left(yz^2 + \left[y\left(x^2 + y\right) + x^2\right] z + x^2\left(x^2 + y\right)\right) dz$$

$$= \frac{1}{x^2 + \frac{1}{2} y} \left[y \frac{z^3}{3} + \left[y\left(x^2 + y\right) + x^2\right] \frac{z^2}{2} + x^2\left(x^2 + y\right) z \right]_{z=0}^{z=1}$$

$$= \frac{1}{x^2 + \frac{1}{2} y} \left(y \frac{1}{3} + \left[y\left(x^2 + y\right) + x^2\right] \frac{1}{2} + x^2\left(x^2 + y\right) \right)$$

Conditional Expectation of a Function - Example

Thus the conditional expectation of $U = X^2 + Y + Z$ given $X = x$, $Y = y$.

$$E\left[U \mid x, y\right] = \frac{1}{x^2 + \frac{1}{2} y} \left(y \frac{1}{3} + \left[y\left(x^2 + y\right) + x^2\right] \frac{1}{2} + x^2\left(x^2 + y\right) \right)$$

$$= \frac{1}{x^2 + \frac{1}{2} y} \left(\frac{y}{3} + \frac{x^2}{2} + \left(x^2 + \frac{1}{2} y\right)\left(x^2 + y\right) \right)$$

$$= \frac{\frac{1}{2} x^2 + \frac{1}{3} y}{x^2 + \frac{1}{2} y} + x^2 + y$$

A Useful Tool: Iterated Expectations

Theorem

Let $(x_1, x_2, \ldots, x_q, y_1, y_2, \ldots, y_m) = (\mathbf{x}, \mathbf{y})$ denote $q + m$ RVs.

Let $U(x_1, x_2, \ldots, x_q, y_1, y_2, \ldots, y_m) = g(\mathbf{x}, \mathbf{y})$. Then,

$$E[U] = E_{\mathbf{y}}\left[E[U|\mathbf{y}]\right]$$

$$Var[U] = E_{\mathbf{y}}\left[Var[U|\mathbf{y}]\right] + Var_{\mathbf{y}}\left[E[U|\mathbf{y}]\right]$$

The first result is commonly referred as the *Law of iterated expectations*. It relates the conditional mean to the unconditional mean.

The second result is referred as the *Law of total variance* or *variance decomposition formula*. It decomposes the variance into a conditional expectation of the conditional variance and a conditional variance of the conditional expectation (variance of regression).

A Useful Tool: Iterated Expectations

Proof: (in the simple case of 2 variables X and Y)

First, we prove the Law of iterated expectations.

Thus $U = g(X, Y)$

$$E[U] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy$$

$$E[U|Y] = E[g(X, Y)|Y] = \int_{-\infty}^{\infty} g(x, y) f_{X|Y}(x|y)\, dx$$

$$= \int_{-\infty}^{\infty} g(x, y) \frac{f(x, y)}{f_Y(y)}\, dx$$

hence $$E_Y\left[E[U|Y]\right] = \int_{-\infty}^{\infty} E[U|y] f_Y(y)\, dy$$

A Useful Tool: Iterated Expectations

$$E_Y\left[E\left[U|Y\right]\right]=\int_{-\infty}^{\infty}E\left[U|y\right]f_Y\left(y\right)dy$$

$$=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}g\left(x,y\right)\frac{f\left(x,y\right)}{f_Y\left(y\right)}dx\right]f_Y\left(y\right)dy$$

$$=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}g\left(x,y\right)f\left(x,y\right)dx\right]dy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g\left(x,y\right)f\left(x,y\right)dxdy=E\left[U\right]$$

A Useful Tool: Iterated Expectations

Now, for the Law of total variance:

$$Var\left[U\right]=E\left[U^2\right]-\left(E\left[U\right]\right)^2$$

$$=E_Y\left[E\left[U^2|Y\right]\right]-\left(E_Y\left[E\left[U|Y\right]\right]\right)^2$$

$$=E_Y\left[Var\left[U|Y\right]+\left(E\left[U|Y\right]\right)^2\right]-\left(E_Y\left[E\left[U|Y\right]\right]\right)^2$$

$$=E_Y\left[Var\left[U|Y\right]\right]+E_Y\left[\left(E\left[U|Y\right]\right)^2\right]-\left(E_Y\left[E\left[U|Y\right]\right]\right)^2$$

$$=E_Y\left[Var\left[U|Y\right]\right]+Var_Y\left(E\left[U|Y\right]\right)$$

A Useful Tool: Iterated Expectations - Example

Example:

Suppose that a rectangle is constructed by first choosing its length, X and then choosing its width Y.

Its length X is selected form an exponential distribution with mean $\mu = {}^1/_\lambda = 5$. Once the length has been chosen its width, Y, is selected from a uniform distribution form 0 to half its length.

Find the mean and variance of the area of the rectangle $A = XY$.

A Useful Tool: Iterated Expectations - Example

Solution:

$$f_X(x) = \tfrac{1}{5} e^{-\frac{1}{5}x} \quad \text{for } \ x \geq 0$$

$$f_{Y|X}(y|x) = \frac{1}{x/2} \quad \text{if } \ 0 \leq y \leq x/2$$

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

$$= \tfrac{1}{5} e^{-\frac{1}{5}x} \frac{1}{x/2} = \tfrac{2}{5x} e^{-\frac{1}{5}x} \quad \text{if } \ 0 \leq y \leq x/2,\, x \geq 0$$

We could compute the mean and variance of $A = XY$ from the joint density $f(x,y)$

A Useful Tool: Iterated Expectations - Example

$$E[A] = E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dxdy$$

$$= \int_0^{\infty}\int_0^{x/2} xy\,\tfrac{2}{5x}e^{-\frac{1}{5}x}dydx = \tfrac{2}{5}\int_0^{\infty}\int_0^{x/2} ye^{-\frac{1}{5}x}dydx$$

$$E\left[A^2\right] = E\left[X^2Y^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2y^2 f(x,y)\,dxdy$$

$$= \int_0^{\infty}\int_0^{x/2} x^2y^2\,\tfrac{2}{5x}e^{-\frac{1}{5}x}dydx = \tfrac{2}{5}\int_0^{\infty}\int_0^{x/2} xy^2e^{-\frac{1}{5}x}dydx$$

and $Var(A) = E\left[A^2\right] - \left(E[A]\right)^2$

A Useful Tool: Iterated Expectations - Example

$$E[A] = \tfrac{2}{5}\int_0^{\infty}\int_0^{x/2} ye^{-\frac{1}{5}x}dydx = \tfrac{2}{5}\int_0^{\infty} e^{-\frac{1}{5}x}\left[\frac{y^2}{2}\right]_{y=0}^{y=x/2} dx$$

$$= \tfrac{2}{5}\tfrac{1}{8}\int_0^{\infty} x^2e^{-\frac{1}{5}x}dx = \tfrac{1}{20}\frac{\Gamma(3)}{\left(\frac{1}{5}\right)^3}\int_0^{\infty}\frac{\left(\frac{1}{5}\right)^3}{\Gamma(3)}x^2e^{-\frac{1}{5}x}dx$$

$$= \tfrac{1}{20}\frac{\Gamma(3)}{\left(\frac{1}{5}\right)^3} = \tfrac{5^3}{20}2 = \tfrac{125}{10} = \tfrac{25}{2} = 12.5$$

A Useful Tool: Iterated Expectations - Example

$$E\left[A^2\right] = \tfrac{2}{5}\int_0^{\infty}\int_0^{x/2} xy^2 e^{-\frac{1}{5}x}\,dy\,dx = \tfrac{2}{5}\int_0^{\infty} xe^{-\frac{1}{5}x}\left[\frac{y^3}{3}\right]_{y=0}^{y=x/2} dx$$

$$= \tfrac{2}{5}\tfrac{1}{3}\tfrac{1}{8}\int_0^{\infty} x^4 e^{-\frac{1}{5}x}\,dx = \tfrac{1}{60}\frac{\Gamma(5)}{\left(\frac{1}{5}\right)^5}\int_0^{\infty}\frac{\left(\frac{1}{5}\right)^5}{\Gamma(5)} x^4 e^{-\frac{1}{5}x}\,dx$$

$$= \tfrac{1}{60}\frac{\Gamma(5)}{\left(\frac{1}{5}\right)^5} = \tfrac{5^5}{60}4! = \tfrac{5^4}{12}24 = 5^4 \times 2 = 1250$$

Thus $Var(A) = E\left[A^2\right] - \left(E[A]\right)^2$

$$= 1250 - (12.5)^2 = 1093.75$$

A Useful Tool: Iterated Expectations - Example

Now, let's use the previous theorem. That is,

$$E[A] = E[XY] = E_X\left[E\left[XY|X\right]\right]$$

and $\quad Var[A] = Var[XY]$

$$= E_X\left[Var\left[XY|X\right]\right] + Var_X\left[E\left[XY|X\right]\right]$$

Now $\quad E\left[XY|X\right] = XE\left[Y|X\right] = X\frac{X}{4} = \tfrac{1}{4}X^2$

and $\quad Var\left(XY|X\right) = X^2 Var\left[Y|X\right] = X^2\frac{\left(X/2-0\right)^2}{12} = \tfrac{1}{48}X^4$

This is because given *X*, *Y* has a uniform distribution from 0 to *X*/2

A Useful Tool: Iterated Expectations - Example

Thus $E[A] = E[XY] = E_X\left[E[XY|X]\right]$

$$= E_X\left[\tfrac{1}{4}X^2\right] = \tfrac{1}{4}E_X\left[X^2\right] = \tfrac{1}{4}\mu_2$$

where $\mu_2 = 2^{nd}$ moment for the exponential dist'n with $\lambda = \frac{1}{5}$

Note $\mu_k = \dfrac{k!}{\lambda^k}$ for the exponential distn

Thus $E[A] = \tfrac{1}{4}\mu_2 = \tfrac{1}{4}\dfrac{2}{\left(\frac{1}{5}\right)^2} = \dfrac{25}{2} = 12.5$

Note: Same answer as previously calculated. No integration needed.

A Useful Tool: Iterated Expectations - Example

Now $E[XY|X] = \tfrac{1}{4}X^2$ and $Var(XY|X) = \tfrac{1}{48}X^4$

Also $Var[A] = Var[XY]$

$$= E_X\left[Var[XY|X]\right] + Var_X\left[E[XY|X]\right]$$

$$E_X\left[Var[XY|X]\right] = E_X\left[\tfrac{1}{48}X^4\right] = \tfrac{1}{48}\mu_4 = \tfrac{1}{48}\frac{4!}{\left(\frac{1}{5}\right)^4} = \frac{5^4}{2}$$

$$Var_X\left[E[XY|X]\right] = Var_X\left[\tfrac{1}{4}X^2\right] = \left(\tfrac{1}{4}\right)^2 Var_X\left[X^2\right]$$

$$= \left(\tfrac{1}{4}\right)^2\left[E_X\left[X^4\right] - \left(E_X\left[X^2\right]\right)^2\right] = \left(\tfrac{1}{4}\right)^2\left[\mu_4 - \left(\mu_2\right)^2\right]$$

A Useful Tool: Iterated Expectations - Example

$$Var_X\left[E\left[XY|X\right]\right] = Var_X\left[\tfrac{1}{4}X^2\right] = \left(\tfrac{1}{4}\right)^2 Var_X\left[X^2\right]$$

$$= \left(\tfrac{1}{4}\right)^2\left[\frac{4!}{\left(\frac{1}{5}\right)^4} - \left(\frac{2!}{\left(\frac{1}{5}\right)^2}\right)^2\right] = \tfrac{5^4}{4^2}\left[4! - (2!)^2\right] = \tfrac{5^4}{4^2}20 = \tfrac{5^5}{4}$$

Thus $\quad Var[A] = Var[XY]$

$$= E_X\left[Var\left[XY|X\right]\right] + Var_X\left[E\left[XY|X\right]\right]$$

$$= \frac{5^4}{2} + \frac{5^5}{4} = 5^4\left(\frac{1}{2} + \frac{5}{4}\right) = 5^4\left(\frac{14}{8}\right) = 1093.75$$

• The same answer as previously calculated!! And no integration needed!

The Multivariate MGF

Definition: **Multivariate MGF**

Let $X_1, X_2, \ldots, X_q$ be q random variables with a joint density function given by $f(x_1, x_2, \ldots, x_q)$. The multivariate MGF is

$$m_{\mathbf{X}}(\mathbf{t}) = E_{\mathbf{X}}[\exp(\mathbf{t}'\mathbf{X})]$$

where **t'**= $(t_1, t_2, \ldots, t_q)$ and **X=** $(X_1, X_2, \ldots, X_q)'$.

If $X_1, X_2, \ldots, X_n$ are n independent random variables, then

$$m_{\mathbf{X}}(\mathbf{t}) = \prod_{i=1}^{n} m_{X_i}(t_i)$$

The MGF of a Multivariate Normal

Definition: **MGF for the Multivariate Normal**

Let $X_1, X_2, \dots, X_q$ be n normal random variables. The multivariate normal MGF is

$$m_{\mathbf{X}}(\mathbf{t}) = E_{\mathbf{X}}[\exp(\mathbf{t}'\mathbf{X})] = \exp(\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\Sigma\,\mathbf{t})$$

where $\mathbf{t}= (t_1, t_2, \dots, t_q)'$, $\mathbf{X}= (X_1, X_2, \dots, X_q)'$ and $\boldsymbol{\mu}= (\mu_1, \mu_2, \dots, \mu_q)'$.

Review: The Transformation Method

Theorem

Let X denote a random variable with probability density function $f(x)$ and $U = h(X)$.

Assume that $h(x)$ is either strictly increasing (or decreasing) then the probability density of U is:

$$g(u) = f\left(h^{-1}(u)\right)\left|\frac{dh^{-1}(u)}{du}\right| = f(x)\left|\frac{dx}{du}\right|$$

The Transformation Method (many variables)

Theorem

Let $x_1, x_2, \ldots, x_n$ denote random variables with joint probability density function

$f(x_1, x_2, \ldots, x_n)$

Let $u_1 = h_1(x_1, x_2, \ldots, x_n)$.

$u_2 = h_2(x_1, x_2, \ldots, x_n)$.

$u_n = h_n(x_1, x_2, \ldots, x_n)$.

define an invertible transformation from the x's to the u's

The Transformation Method (many variables)

Then the joint probability density function of $u_1, u_2, \ldots, u_n$ is given by:

$$g(u_1, \cdots, u_n) = f(x_1, \cdots, x_n)\left|\frac{d(x_1, \cdots, x_n)}{d(u_1, \cdots, u_n)}\right|$$

$$= f(x_1, \cdots, x_n)|J|$$

where $J = \dfrac{d(x_1, \cdots, x_n)}{d(u_1, \cdots, u_n)} = \det\begin{bmatrix} \dfrac{dx_1}{du_1} & \cdots & \dfrac{dx_1}{du_n} \\ \vdots & & \vdots \\ \dfrac{dx_n}{du_1} & \cdots & \dfrac{dx_n}{du_n} \end{bmatrix}$

Jacobian of the transformation

Example: Distribution of x+y and x-y

Suppose that x_1, x_2 are independent with density functions $f_1(x_1)$ and $f_2(x_2)$

Find the distribution of $u_1 = x_1 + x_2$ and $u_2 = x_1 - x_2$

Solution: Solving for x_1 and x_2, we get the inverse transformation:

$$x_1 = \frac{u_1 + u_2}{2} \qquad x_2 = \frac{u_1 - u_2}{2}$$

The Jacobian of the transformation

$$J = \frac{d(x_1, x_2)}{d(u_1, u_2)} = \det \begin{bmatrix} \frac{dx_1}{du_1} & \frac{dx_1}{du_2} \\ \frac{dx_2}{du_1} & \frac{dx_2}{du_2} \end{bmatrix}$$

Example: Distribution of x+y and x-y

$$J = \frac{d(x_1, x_2)}{d(u_1, u_2)} = \det \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right) - \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = -\frac{1}{2}$$

The joint density of x_1, x_2 is

$$f(x_1, x_2) = f_1(x_1)\, f_2(x_2)$$

Hence the joint density of u_1 and u_2 is:

$$g(u_1, u_2) = f(x_1, x_2)\,|J|$$

$$= f_1\left(\frac{u_1 + u_2}{2}\right) f_2\left(\frac{u_1 - u_2}{2}\right) \frac{1}{2}$$

Example: Distribution of x+y and x-y

From $$g(u_1,u_2) = f_1\left(\frac{u_1+u_2}{2}\right) f_2\left(\frac{u_1-u_2}{2}\right)\frac{1}{2}$$

We can determine the distribution of $u_1 = x_1 + x_2$

$$g_1(u_1) = \int_{-\infty}^{\infty} g(u_1,u_2)\,du_2$$

$$= \int_{-\infty}^{\infty} f_1\left(\frac{u_1+u_2}{2}\right) f_2\left(\frac{u_1-u_2}{2}\right)\frac{1}{2}\,du_2$$

$$\text{put } v = \frac{u_1+u_2}{2} \text{ then } \frac{u_1-u_2}{2} = u_1 - v, \frac{dv}{du_2} = \frac{1}{2}$$

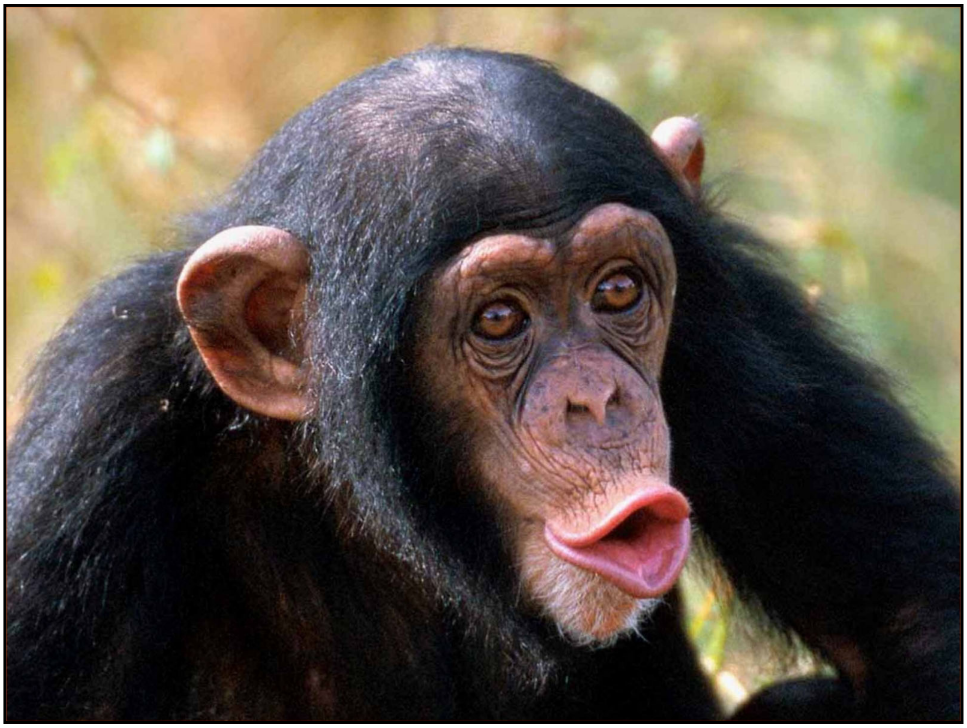
Example: Distribution of x+y and x-y

Hence

$$g_1(u_1) = \int_{-\infty}^{\infty} f_1\left(\frac{u_1+u_2}{2}\right) f_2\left(\frac{u_1-u_2}{2}\right)\frac{1}{2}\,du_2$$

$$= \int_{-\infty}^{\infty} f_1(v)\, f_2(u_1 - v)\,dv$$

This is called the *convolution* of the two densities f_1 and f_2.



Example (1): Convolution formula -The Gamma distribution

Let X and Y be two independent random variables such that X and Y have an exponential distribution with parameter λ.

We will use the convolution formula to find the distribution of $U = X + Y$. (We already know the distribution of U: Gamma.)

$$g_U(u) = \int_{-\infty}^{\infty} f_U(u-y) f_Y(y) dy = \int_0^u \lambda\, e^{-\lambda(u-y)} \lambda\, e^{-\lambda y} dy$$

$$= \int_0^u \lambda^2\, e^{-\lambda u} dy = \lambda^2\, u e^{-\lambda u}$$

This is the gamma distribution when $\alpha=2$.

Example (2): The ex-Gaussian distribution

Let X and Y be two independent random variables such that:

1. X has an exponential distribution with parameter λ.
2. Y has a normal (Gaussian) distribution with mean μ and standard deviation σ.

We will use the convolution formula to find the distribution of $U = X + Y$.

(This distribution is used in psychology as a model for response time to perform a task.)

Example (2): The ex-Gaussian distribution

Now
$$f_1(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$f_2(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The density of $U = X + Y$ is:

$$g(u) = \int_{-\infty}^{\infty} f_1(v)\, f_2(u - v)\, dv$$

$$= \int_{0}^{\infty} \lambda e^{-\lambda v} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(u-v-\mu)^2}{2\sigma^2}} dv$$

Example (2): The ex-Gaussian distribution

or $$g(u) = \frac{\lambda}{\sqrt{2\pi}\sigma}\int_0^\infty e^{-\frac{(u-v-\mu)^2}{2\sigma^2}-\lambda v}\,dv$$

$$= \frac{\lambda}{\sqrt{2\pi}\sigma}\int_0^\infty e^{-\frac{(u-v-\mu)^2+2\sigma^2\lambda v}{2\sigma^2}}\,dv$$

$$= \frac{\lambda}{\sqrt{2\pi}\sigma}\int_0^\infty e^{-\frac{v^2-2(u-\mu)v+(u-\mu)^2+2\sigma^2\lambda v}{2\sigma^2}}\,dv$$

$$= \frac{\lambda}{\sqrt{2\pi}\sigma}e^{-\frac{(u-\mu)^2}{2\sigma^2}}\int_0^\infty e^{-\frac{v^2-2[(u-\mu)-\sigma^2\lambda]v}{2\sigma^2}}\,dv$$

Example (2): The ex-Gaussian distribution

or $$= \frac{\lambda}{\sqrt{2\pi}\sigma}e^{-\frac{(u-\mu)^2}{2\sigma^2}}\int_0^\infty e^{-\frac{v^2-2[(u-\mu)-\sigma^2\lambda]v}{2\sigma^2}}\,dv$$

$$= \frac{\lambda}{\sqrt{2\pi}\sigma}e^{-\frac{(u-\mu)^2-[(u-\mu)-\sigma^2\lambda]^2}{2\sigma^2}}\int_0^\infty e^{-\frac{v^2-2[(u-\mu)-\sigma^2\lambda]v+[(u-\mu)-\sigma^2\lambda]^2}{2\sigma^2}}\,dv$$

$$= \lambda e^{-\frac{(u-\mu)^2-[(u-\mu)-\sigma^2\lambda]^2}{2\sigma^2}}\int_0^\infty \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{v^2-2[(u-\mu)-\sigma^2\lambda]v+[(u-\mu)-\sigma^2\lambda]^2}{2\sigma^2}}\,dv$$

$$= \lambda e^{-\frac{(u-\mu)^2-[(u-\mu)-\sigma^2\lambda]^2}{2\sigma^2}}\,P[V \geq 0]$$

Example (2): The ex-Gaussian distribution

Where V has a Normal distribution with mean

$$\mu_V = u - \left(\mu + \sigma^2 \lambda\right)$$

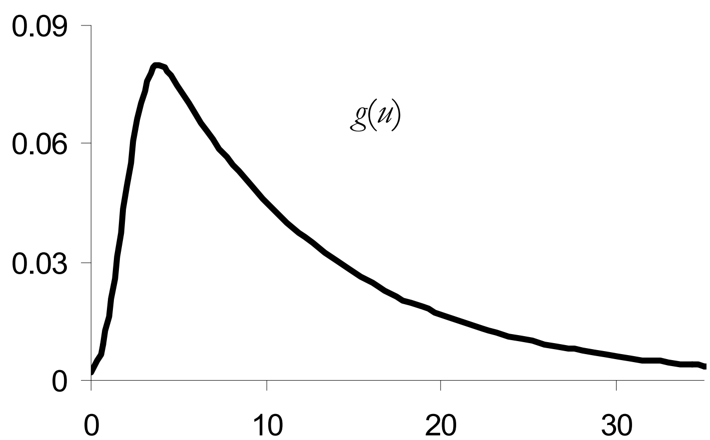
and variance σ^2.

That is,

$$g(u) = \lambda e^{-\lambda\left[(u-\mu)-\frac{\sigma^2\lambda}{2}\right]}\left[1 - \Phi\left(\frac{\left[\mu + \sigma^2\lambda\right] - u}{\sigma^2}\right)\right]$$

Where $\Phi(z)$ is the cdf of the standard Normal distribution

The ex-Gaussian distribution



Distribution of Quadratic Forms

We will present different theorems when the RVs are normal variables:

Theorem 7.1. If $\mathbf{y} \sim N(\mu_y, \mathbf{\Sigma}_y)$, then $\mathbf{z} = \mathbf{Ay} \sim N(\mathbf{A}\mu_y, \mathbf{A}\,\mathbf{\Sigma}_y\,\mathbf{A}')$, where $\mathbf{A}$ is a matrix of constants and $\mathbf{y}$ a $n\times 1$ vector .

Theorem 7.2. Let $\mathbf{y}$ be a $n\times 1$ vector $\sim N(0, \mathbf{I}_n)$. Then $\mathbf{y'y} \sim \chi_n^2$.

Theorem 7.3. Let the $n\times 1$ vector $\boldsymbol{y} \sim N(\mathbf{0}, \sigma^2\, \mathbf{I}_n)$ and $\mathbf{M}$ be a symmetric idempotent matrix of rank m. Then,

$\mathbf{y}'\mathrm{M}\mathbf{y}/\sigma^2 \sim \chi_{tr(M)}^2$.

Proof: Since $\mathbf{M}$ is symmetric it can be diagonalized with an orthogonal matrix $\mathbf{Q}$. That is, $\mathbf{Q'MQ} = \mathbf{\Lambda}$. $(Q'Q=I)$

Since $\mathbf{M}$ is idempotent all these roots are either 0 or 1. Thus,

$$Q'MQ = \Lambda = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

Note: $\dim(\mathbf{I}) = \text{rank}(\mathbf{M})$ (the number of non-zero roots is the rank of the matrix). Also, since $\Sigma_i \lambda_i = \text{tr}(\mathbf{I})$, $\Rightarrow \dim(\mathbf{I}) = \text{tr}(\mathbf{M})$.

Let $v = \mathbf{Q}'\boldsymbol{y}$.

$E(v) = \mathbf{Q}'E(\boldsymbol{y}) = 0$

$\text{Var}(v) = E[vv'] = E[\mathbf{Q}'\boldsymbol{yy}'\mathbf{Q}] = \mathbf{Q}'E(\sigma^2\mathbf{I}_n)\mathbf{Q} = \sigma^2\, \mathbf{Q}'\mathbf{I}_n\mathbf{Q} = \sigma^2\, \mathbf{I}_n$

$\Rightarrow v \sim N(0, \sigma^2 \mathbf{I}_n)$

Then,
$$\frac{y'My}{\sigma^2} = \frac{v'Q'MQv}{\sigma^2} = \frac{1}{\sigma^2} v' \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} v = \frac{1}{\sigma^2} \sum_{i=1}^{tr(M)} v_i^2 = \sum_{i=1}^{tr(M)} \left(\frac{v_i}{\sigma}\right)^2$$

Thus, $\mathbf{y}'\mathbf{My}/\sigma^2$ is the sum of $\text{tr}(\mathbf{M})$ $N(0,1)$ squared variables. It follows a $\chi_{tr(M)}^2$.

Theorem 7.4. Let the $n\times1$ vector $\boldsymbol{y} \sim N(\mu_y, \Sigma_y)$. Then,

$(\boldsymbol{y} - \mu_y)' \Sigma_y^{-1} (\boldsymbol{y} - \mu_y) \sim \chi_n^2$.

Proof:

Recall that there exists a non-singular matrix **A** such that $\mathbf{AA}' = \Sigma_y$.

Let $v = \mathbf{A}^{-1} (\boldsymbol{y} - \mu_y)'$ (a linear combination of normal variables)

$\Rightarrow v \sim N(0, \mathbf{I}_n)$

$\Rightarrow v' v \sim \chi_n^2$. (using Theorem 7.3, where $n = \text{tr}(\Sigma_y^{-1})$.)

Theorem 7.5

Let the $n\times1$ vector $\boldsymbol{y} \sim N(0, \mathbf{I})$ and **M** be an $n\times n$ matrix. Then, the characteristic function of $\boldsymbol{y}'\mathbf{M}\boldsymbol{y}$ is $|\mathbf{I}-2it\mathbf{M}|^{-1/2}$

Proof:

$$\varphi_{y'My} = E_y[e^{ity'My}] = \frac{1}{(2\pi)^{n/2}} \int_y e^{ity'My} e^{-y'y/2} dx = \frac{1}{(2\pi)^{n/2}} \int_y e^{-y'(I-2itM)y/2} dx.$$

This is the normal density with $\Sigma^{-1} = (\mathbf{I}-2it\mathbf{M})$, except for the determinant $|\mathbf{I}-2it\mathbf{M}|^{-1/2}$, which should be in the denominator.

Theorem 7.6

Let the $n\times1$ vector $\boldsymbol{y} \sim N(0, I)$, **M** be an $n\times n$ idempotent matrix of rank m, let **L** be an $n\times n$ idempotent matrix of rank s, and suppose $\mathbf{ML} = 0$. Then, $\mathbf{y'My}$ and $\mathbf{y'Ly}$ are independently distributed χ^2 variables.

Proof:

By Theorem 7.3 both quadratic forms χ^2 distributed variables. We only need to prove independence. From Theorem 7.5, we have

$$\varphi_{y'My} = E_y[e^{ity'My}] = |I-2itM|^{-1/2}$$

$$\varphi_{y'Ly} = E_y[e^{ity'Ly}] = |I-2itL|^{-1/2}$$

The forms will be independently distributed if $\varphi_{y'(M+L)y} = \varphi_{y'My}\, \varphi_{y'Ly}$

That is,

$$\varphi_{y'(M+L)y} = E_y[e^{ity'(M+L)y}] = |I-2it(M+L)|^{-1/2} = |I-2itM|^{-1/2} |I-2itL|^{-1/2}$$

Since $|\mathbf{ML}| = |\mathbf{M}|\,|\mathbf{L}|$, the result will be true only when $\mathbf{ML}=0$.