# Fundamental Tools - Probability Theory III

MSc Financial Mathematics

The University of Warwick

September 30, 2015

# Joint distribution function

For two random variables $X$ and $Y$, we define their joint distribution function by

$$F_{XY}(x,y) = \mathbb{P}(X \leqslant x, Y \leqslant y), \quad -\infty < x, y < \infty.$$

Two important classes:

- $X$ and $Y$ are jointly discrete if both $X$ and $Y$ are discrete. It is convenient to work with their joint probability mass function

$$p_{XY}(x,y) = \mathbb{P}(X = x, Y = y).$$

- $X$ and $Y$ are jointly continuous if there exists a non-negative function $f_{XY} : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\mathbb{P}(X \leqslant x, Y \leqslant y) = F_{XY}(x,y) = \int_{u=-\infty}^{u=x} \int_{v=-\infty}^{v=y} f_{XY}(u,v) dv du.$$

  We call $f_{XY}$ the joint probability density function of $X$ and $Y$.

# Recovering marginal distribution

Given a joint probability distribution/mass/density function of $(X, Y)$, we can recover the the corresponding marginal characteristics of $X$ as follows:

- Marginal distribution function

$$F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X \leqslant x, Y < \infty) = F_{XY}(x, \infty).$$

- Marginal probability mass function (if $(X, Y)$ are jointly discrete)

$$p_X(x) = \mathbb{P}(X = x) = \mathbb{P}(X = x, Y \in S_Y) = \sum_y p_{XY}(x, y).$$

- Marginal probability density function (if $(X, Y)$ are jointly continuous)

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} F_{XY}(x, \infty) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy.$$

# Working with jointly continuous random variables

From definition, the joint distribution and density function of $(X, Y)$ are related via

$$f_{XY}(x,y) = \frac{d^2}{dxdy} F_{XY}(x,y).$$

In univariate case, we compute probabilities involving a continuous random variable via simple integration:

$$\mathbb{P}(X \in A) = \int_A f(x)dx.$$

In bivariate case, probabilities are computed via double integration

$$\mathbb{P}((X,Y) \in A) = \iint_A f_{XY}(x,y)dxdy.$$

and the calculation is not necessarily straightforward.

# Example 1

Let $(X, Y)$ be a pair of jointly continuous random variables with joint density function $f(x, y) = 1$ on $(x, y) \in [0, 1]^2$ (and is zero elsewhere). Find $\mathbb{P}(X - Y < 1/2)$.

*Sketch of solution*
Define the set $A = \{(x, y) : (x, y) \in [0, 1]^2, x - y \leqslant 1/2\}$. Then

$$\mathbb{P}(X - Y < 1/2) = \mathbb{P}((X, Y) \in A) = \iint_A f(x, y) dxdy = \iint_A dxdy$$
$$= \text{The area of } A.$$

By sketching the set $A$, the area is found to be $7/8$.

# Example 2

Let $(X, Y)$ be a pair of jointly continuous random variables with joint density function $f(x, y) = e^{-y}$ on $0 < x < y < \infty$ (and is zero elsewhere). Verify that the given $f$ is a well-defined joint density function. Find the marginal density function of $X$.

*Sketch of solution*

To show that $f_{XY}$ is a well-defined joint density function, one needs to show $f \geqslant 0$ and $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy = 1$. The first property is obvious. To show the latter, define $A = \{(x, y) : 0 < x < y < \infty\}$. Then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy = \iint_A e^{-y}dxdy = \int_0^{\infty}\int_0^{y} e^{-y}dxdy$$
$$= \int_0^{\infty} ye^{-y}dy = 1.$$

The marginal density of $X$ is, for $x > 0$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_x^{\infty} e^{-y}dy = e^{-y}|_{\infty}^{x} = e^{-x}.$$

# Expected value involving joint distribution

Let $(X, Y)$ be a pair of random variables. For a given function $g(\cdot,\cdot)$, the expected value of the random variable $g(X, Y)$ is given by

$$\mathbb{E}(g(X,Y)) = \begin{cases} \sum_{x,y} g(x,y) p_{XY}(x,y); \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y) dxdy. \end{cases}$$

Two important specifications of $g(\cdot,\cdot)$:

- Set $g(x, y) = x + y$. One could obtain $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.
- Set $g(x, y) = xy$. This leads to computation of the covariance measure between $X$ and $Y$ defined by

  $$Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

  and correlation measure defined by

  $$corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{\text{var}(X)\,\text{var}(Y)}}.$$

# Conditional distributions

If $(X, Y)$ are jointly discrete, the conditional probability mass function of $X$ given $Y = y$ is

$$p_{X|Y}(x|y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{p_{XY}(x, y)}{p_Y(y)}.$$

If $(X, Y)$ are jointly continuous, the conditional probability density function of $X$ given $Y = y$ is

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}$$

which can be interpreted as follows:

$$\begin{aligned}\mathbb{P}(x \leqslant X \leqslant x + \delta x | y \leqslant Y \leqslant y + \delta y) &= \frac{\mathbb{P}(x \leqslant X \leqslant x + \delta x, y \leqslant Y \leqslant y + \delta y)}{\mathbb{P}(y \leqslant Y \leqslant y + \delta y)} \\ &\approx \frac{f_{XY}(x, y)\delta x \delta y}{f_Y(y)\delta y} = f_{X|Y}(x|y)\delta x.\end{aligned}$$

# Conditional probability and expectation

With conditional probability mass/density function, we can work out the conditional probability and expectation as follows:

- Conditional probability:

$$\mathbb{P}(X \in A | Y = y) = \begin{cases} \sum_{x \in A} p_{X|Y}(x|y); \\ \int_A f_{X|Y}(x|y)dx. \end{cases}$$

- Conditional expectation:

$$\mathbb{E}(X|Y = y) = \begin{cases} \sum_x x p_{X|Y}(x|y); \\ \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx. \end{cases}$$

# Example

Let the joint density function of $X$ and $Y$ be $f_{XY}(x,y) = \frac{e^{-x/y}e^{-y}}{y}$ on $0 < x, y < \infty$ (and zero elsewhere). Find the conditional density $f_{X|Y}$, and compute $\mathbb{P}(X > 1|Y = y)$ and $\mathbb{E}(X|Y = y)$.

*Sketch of solution*

We first work out $f_Y$ by integrating $x$ out from $f_{XY}$:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx = \int_0^{\infty} \frac{e^{-x/y}e^{-y}}{y}dx = e^{-y}$$

for $y > 0$. Then $f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{e^{-x/y}}{y}$ for $x > 0$.

$$\mathbb{P}(X > 1|Y = y) = \int_1^{\infty} f_{X|Y}(x|y)dx = \int_1^{\infty} \frac{e^{-x/y}}{y}dx = e^{-1/y}$$

and

$$\mathbb{E}(X|Y = y) = \int_0^{\infty} x f_{X|Y}(x|y)dx = \int_0^{\infty} \frac{xe^{-x/y}}{y}dx = y$$

# Independent random variables

We say that $X$ and $Y$ are independent random variables if

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(X \in B)$$

for any Borel set $A, B \in \mathcal{B}(\mathbb{R})$.

The following are equivalent conditions for $X$ and $Y$ being independent:

- $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$ for all functions $f, g$.
- $p_{XY}(x, y) = p_X(x)p_Y(y)$ or equivalently $p_{X|Y}(x|y) = p_X(x)$ in case $(X, Y)$ are jointly discrete.
- $f_{XY}(x, y) = f_X(x)f_Y(y)$ or equivalently $f_{X|Y}(x|y) = f_X(x)$ in case $(X, Y)$ are jointly continuous.

# Independence and zero correlation

If $X$ and $Y$ are independent, then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. This leads to:

1. $Cov(X, Y) = 0$, which also implies the correlation between $X$ and $Y$ is zero.
2. $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$.

The reverse is not true in general. The most importantly, zero correlation does not imply independence. See problem sheet.

# Sum of independent random variables

We are interested in the following question: Suppose $X$ and $Y$ are two independent random variables. What is the distribution of $X + Y$ then?

The procedure is easier in the discrete case

*Example*: Suppose $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$. Then for $k = 0, 1, 2...$

$$\begin{aligned} p_{X+Y}(k) &= \mathbb{P}(X+Y=k) = \sum_{j=0}^{k} \mathbb{P}(X=j)\mathbb{P}(Y=k-j) \\ &= \sum_{j=0}^{k} \frac{e^{-\lambda_1}\lambda_1^j}{j!}\frac{e^{-\lambda_2}\lambda_2^{k-j}}{(k-j)!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{j=0}^{k} C_j^k \lambda_1^j \lambda_2^{k-j} = \frac{e^{-(\lambda_1+\lambda_2)}}{k!}(\lambda_1+\lambda_2)^k. \end{aligned}$$

This is the pmf of a $Poi(\lambda_1 + \lambda_2)$ random variable.

# Sum of independent random variables

Assume $(X, Y)$ are jointly continuous and let $Z = X + Y$. Then the CDF of $Z$ is

$$F_Z(z) = \mathbb{P}(Z \leqslant z) = \mathbb{P}(X + Y \leqslant z) = \iint_{x+y\leqslant z} f_X(x) f_Y(y) dx dy$$

$$= \int_{y=-\infty}^{y=\infty} \left( \int_{x=-\infty}^{x=z-y} f_X(x) dx \right) f_Y(y) dy$$

$$= \int_{y=-\infty}^{y=\infty} F_X(z-y) f_Y(y) dy.$$

Differentiation w.r.t $z$ gives the density of $Z$ as

$$f_Z(z) = \int_{y=-\infty}^{y=\infty} f_X(z-y) f_Y(y) dy.$$

# Example

Let $X$ and $Y$ be two independent $U[0,1]$ random variables. Find the density function of $Z = X + Y$.

*Sketch of solution*
We have $f_X(x) = 1$ on $x \in [0,1]$ and $f_Y(y) = 1$ on $y \in [0,1]$. Obviously $Z$ can only take value in $[0,2]$. The density of $Z$ is given by

$$f_Z(z) = \int_{y=-\infty}^{y=\infty} f_X(z-y) f_Y(y) dy = \int_A dy$$

where $A = \{y : 0 \leqslant y \leqslant 1, z - 1 \leqslant y \leqslant z\}$. Thus:

- When $0 \leqslant z \leqslant 1$, $A = [0, z]$, and then $f_Z(z) = \int_0^z dy = z$.
- When $1 < z \leqslant 2$, $A = [z-1, 1]$, and then $f_Z(z) = \int_{1-z}^1 dy = 2 - z$.
- And the density function is *zero* elsewhere.

# Probability generating function

Probability generating function is only defined for a discrete random variable $X$ taking values in non-negative integers $\{0, 1, 2, ...\}$. It is defined as

$$G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^{\infty} s^k p_X(k).$$

- View $G_X$ as a Taylor's expansion in $s$:

  $$G_X(s) = p_X(0) + p_X(1)s + p_X(2)s^2 + \cdots$$

  We could then deduce $p_X(n) = \frac{G_X^{(n)}(0)}{n!}$, i.e. $G_X$ uniquely determines the pmf of $X$. In other words, if the probability generating functions of $X$ and $Y$ are equal, then $X$ and $Y$ have the same distribution.

- If $X$ and $Y$ are independent,

  $$G_{X+Y}(s) = \mathbb{E}(s^X s^Y) = \mathbb{E}(s^X)\mathbb{E}(s^Y) = G_X(s)G_Y(s).$$

  Hence we can study the distribution of $X + Y$ via $G_X(s)G_Y(s)$.

# Moments calculation from probability generating function

Given $G_X$, we can derive the moments of $X$.

$$
\begin{aligned}
G_X^{(1)}(s) &= \mathbb{E}(\frac{d}{ds}s^X) = \mathbb{E}(Xs^{X-1}) \\
&\Longrightarrow \mathbb{E}(X) = G_X^{(1)}(1) \\
G_X^{(2)}(s) &= \mathbb{E}(\frac{d^2}{ds^2}s^X) = \mathbb{E}(X(X-1)s^{X-2}) \\
&\Longrightarrow \mathbb{E}(X(X-1)) = G_X^{(2)}(1) \\
G_X^{(3)}(s) &= \mathbb{E}(\frac{d^3}{ds^3}s^X) = \mathbb{E}(X(X-1)(X-2)s^{X-3}) \\
&\Longrightarrow \mathbb{E}(X(X-1)(X-2)) = G_X^{(3)}(1)
\end{aligned}
$$

# Example

Find the pgf of $Poi(\lambda)$. If $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$, what is the distribution of $X + Y$?

*Sketch of solution*
For $N \sim Poi(\lambda)$,

$$G(s) = \mathbb{E}(s^N) = \sum_{k=0}^{\infty} s^k \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(s\lambda)^k}{k!} = e^{-\lambda} e^{s\lambda} = e^{\lambda(s-1)}.$$

Now if $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$ are independent, the pgf of $G_{X+Y}(s)$ is $G_X(s)G_Y(s)$. Then

$$G_{X+Y}(s) = e^{\lambda_1(s-1)} e^{\lambda_2(s-1)} = e^{(\lambda_1+\lambda_2)(s-1)}$$

which is the pgf of a $Poi(\lambda_1 + \lambda_2)$ distribution. Hence we conclude that $X + Y$ has a $Poi(\lambda_1 + \lambda_2)$ distribution by the unique correspondence between pgf and probability mass function.

# Moment generating function (mgf)

Moment generating function (mgf) can be defined for general random variable via

$$m_X(t) = \mathbb{E}(e^{tX}) = \begin{cases} \sum_x e^{tx} p_X(x), & \text{if } X \text{ is discrete;} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx, & \text{if } X \text{ is continuous.} \end{cases}$$

Consider the $n$-th derivative of $m_X(t)$:

$$m_X^{(n)}(t) = \frac{d^n}{dt^n}\mathbb{E}(e^{tX}) = \mathbb{E}(\frac{d^n}{dt^n} e^{tX}) = \mathbb{E}(X^n e^{tX})$$

from which we obtain $\mathbb{E}(X^n) = m_X^{(n)}(0)$ on letting $t = 0$.

A mgf also uniquely determines the underlying distribution:

- If $X$ and $Y$ have the same mgf, then they must have the same distribution.
- Suppose $X$ and $Y$ are independent, the mgf of $X + Y$ is
$$m_{X+Y}(t) = \mathbb{E}(e^{tX} e^{tY}) = \mathbb{E}(e^{tX})\mathbb{E}(e^{tY}) = m_X(t)m_Y(t).$$
Hence we can study the distribution of $X + Y$ via $m_X(t)m_Y(t)$, just like pgf.

# Example

Find the moment generating function of $N(\mu, \sigma^2)$. If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$, what is the distribution of $X + Y$?

$$\begin{aligned} m_X(t) = \mathbb{E}(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}\right) dx \\ &= \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right). \end{aligned}$$

We obtain the mgf of $X + Y$ using the fact $m_{X+Y}(t) = m_X(t) m_Y(t)$:

$$\begin{aligned} m_{X+Y}(t) &= \exp\left(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2\right) \exp\left(\mu_2 t + \frac{1}{2}\sigma_2^2 t^2\right) \\ &= \exp\left((\mu_1 + \mu_2) t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2) t^2\right) \end{aligned}$$

which is the mgf of $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Thus $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ by the unique correspondence between mgf and probability distribution.