Random Vector

▶ Introduce a random experiment with a sample space \( C \) and
\[ X_1(c) = x_1, X_2(c) = x_2 \]

▶ A random vector \((X_1, X_2)\) has the space
\[ D = \{(x_1, x_2) : x_1 = X_1(c), x_2 = X_2(c), c \in C\} \]

▶ The common notation is \( X = (X_1, X_2)' \).

▶ The probability of any event \( A \) can be defined in terms of the (joint) cumulative distribution function

\[ F_{X_1, X_2}(x_1, x_2) = P[\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}] \]

▶ Easy to show that for any rectangle
\[ P[a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2] = F_{X_1, X_2}(b_1, b_2) - F_{X_1, X_2}(a_1, b_2) \]
\[ - F_{X_1, X_2}(b_1, a_2) + F_{X_1, X_2}(a_1, a_2) \]
A random vector $\mathbf{X}$ is a **discrete random vector** if its space $\mathcal{D}$ is countable.

The **joint probability mass function (pmf)** of $(X_1, X_2)$ is

$$p_{X_1,X_2}(x_1, x_2) = P[X_1 = x_1, X_2 = x_2]$$

The pmf uniquely defines the cdf and is characterized by two properties:

1. $0 \leq p_{X_1,X_2}(x_1, x_2) \leq 1$
2. $\sum \sum_{\mathcal{D}} p_{X_1,X_2}(x_1, x_2) = 1$

For any event $B \in \mathcal{D}$

$$P[(X_1, X_2) \in B] = \sum \sum_{B} p_{X_1,X_2}(x_1, x_2)$$
Marginal distributions

- The **support** of a vector \((X_1, X_2)\) is the set of all \(\{(x_1, x_2) : p(x_1, x_2) > 0\}\).

- Let \(D_{X_1}\) be the support of \(X_1\). Note that

\[
F_{X_1}(x_1) = \sum \sum_{w_1 \leq x_1, -\infty < x_2 < \infty} p_{X_1, X_2}(w_1, x_2) = \sum_{w_1 \leq x_1} \left\{ \sum_{x_2 < \infty} p_{X_1, X_2}(w_1, x_2) \right\}
\]

- By uniqueness of cdfs,

\[
p_{X_1}(x_1) = \sum_{x_2 < \infty} p_{X_1, X_2}(x_1, x_2)
\]
A fair die is rolled twice, $X$ is the larger and $Y$ is the smaller of the two.

Necessarily $X \geq Y$ and the sample space is $\mathcal{C} = \{(1, 1), (2, 1), (2, 2), (3, 1), \ldots, (6, 4), (6, 5), (6, 6)\}$.

Easy to check that $p(x, y) = \frac{2}{36}$ for any $x, y = 1, 2, \ldots, 6, x > y$ and $p(x, y) = \frac{1}{36}$ for any $x = y = 1, 2, \ldots, 6$.

For example, $P(X = 1) = \sum_{y=1}^{6} P(X = 1, Y = y) = \frac{1}{36}$ and $P(X = 2) = \sum_{y=1}^{6} P(X = 2, Y = y) = \frac{1}{12}$.
Continuous bivariate random variables

- A random vector $\mathbf{X} = (X_1, X_2)'$ is of the **continuous** type if its cdf $F_{X_1, X_2}(x_1, x_2)$ is continuous.
- For the most part, we assume that there exists a function $f_{X_1, X_2}(x_1, x_2)$ such that

$$F_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1, X_2}(w_1, w_2) \, dw_1 \, dw_2$$

- The integrand is the **joint probability density function** (pdf) of $X_1, X_2$.
- If we know the cdf,

$$f_{X_1, X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1, X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$
Characterization of pdf and probabilities of events

- First: $f_{X_1,X_2}(x_1, x_2) \geq 0$
- Second: $\int \int_D f_{X_1,X_2}(x_1, x_2) \, dx_1 \, dx_2 = 1$
- For any event $A \in D$ we have
  \[ P[(X_1, X_2) \in A] = \int \int_A f_{X_1,X_2}(x_1, x_2) \, dx_1, \, dx_2 \]
- For a continuous random vector $(X_1, X_2)$ the support of $(X_1, X_2)$ is the set of all points $(x_1, x_2)$ such that $p(x_1, x_2) > 0$
The marginal cdf is

\[ F_{X_1}(x_1) = P[X_1 \leq x_1, -\infty < X_2 < \infty] = \int_{-\infty}^{\infty} F_{X_1,X_2}(x_1, x_2) \, dx_2 \]

In the same way as for discrete variable, the marginal pmf is

\[ f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) \, dx_2 \]
Examples

- **Bivariate Uniform** - \( f(x, y) = 1 \) in the \([0, 1] \times [0, 1]\) rectangle

- Check that the marginal density \( f(x) = 1 \) for any \( 0 \leq x \leq 1 \); thus, both \( X \sim Unif[0, 1] \) and \( Y \sim Unif[0, 1] \)

- Uniform distribution on a triangle: \( f(x, y) = 2 \) for any \( x, y \geq 0 \) and \( x + y \leq 1 \)

- In that case, the marginal density of \( X \) is

\[
f_1(x) = \int_0^{1-x} 2 \, dy = 2(1 - x)
\]

for \( 0 \leq x \leq 1 \)
A joint density $f(x, y) = x \exp(-x(1+y))$ for $x, y \geq 0$

Verify that $f_1(x) = \exp(-x)$ and $f_2(y) = \frac{1}{(1+y)^2}$

Another example: the joint density $f(x, y) = c - 2(c-1)(x + y - 2xy)$ for $x, y \in [0, 1]$

This is a joint density for any constant $c$; check that both $f_1(x) = 1$ and $f_2(y) = 1$ for both $x, y \in [0, 1]$

Thus, both marginals are uniform while the joint density is not
Under the assumption that
\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| g(x_1, x_2) \right| f(x_1, x_2) \, dx_1 \, dx_2 < \infty \]
the expectation of \( Y = g(X_1, X_2) \) is
\[ \mathbb{E} Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) f(x_1, x_2) \, dx_1 \, dx_2 < \infty \]

Very important: \( \mathbb{E} \) is still a linear operator: for any two constants \( k_1, k_2 \)
\[ \mathbb{E}(k_1 Y_1 + k_2 Y_2) = k_1 \mathbb{E} Y_1 + k_2 \mathbb{E} Y_2 \]
Expectation of the function of just one coordinate

To find $E g(X_2)$, can do either of two:

$$E(g(X_2)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_2)f(x_1, x_2)dx_1 dx_2 = \int_{-\infty}^{\infty} g(x_2)f_{X_2}(x_2)dx_2$$
Examples

▶ Bivariate Uniform on $[0, 1]^2$
▶ The expected distance between the two coordinates is

$$
\int_0^1 \int_0^1 |x-y| \, dx \, dy = \int_0^1 \left[ \int_0^y (y-x) \, dx + \int_y^1 (x-y) \, dx \right] \, dy = \frac{1}{3}
$$
Examples

- The uniform density on a triangle $x, y \geq 0$ and $x + y \leq 1$ is $f(x, y) = 2$ - recall a previous example
- The marginal density of $X$ is $f(x) = 2(1 - x)$ for $0 \leq x \leq 1$. Thus,
  \[ \mathbb{E}(X) = \int_0^1 2x(1 - x)dx = \frac{1}{3} \]
- Next,
  \[ \mathbb{E}(X^2) = \int_0^1 2x^2(1 - x)dx = \frac{1}{6} \]
  and so $\text{Var}Y = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$
- Also,
  \[ \mathbb{E}(XY) = 2 \int_0^1 \int_0^{1-y} xy dxdy = \frac{1}{12} \]
  Therefore, $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = -\frac{1}{36}$ and $\text{Corr}(X, Y) = -\frac{1}{2}$
A moment generating function of a random vector $X$ is

$$M_{X_1, X_2}(t_1, t_2) = E \exp (t' X)$$

We need positive $h_1$ and $h_2$ such that $M_{X_1, X_2}(t_1, t_2)$ is finite for $|t_1| < h_1$ and $|t_2| < h_2$

$$M_{X_1, X_2}(t_1, 0) = M_{X_1}(t_1) \text{ and } M_{X_1, X_2}(0, t_2) = M_{X_2}(t_2)$$
Example

- Let \((X, Y)\) have the joint pdf

\[ f(x, y) = \exp(-y) \]

for any \(0 < x < y < \infty\)

- The mgf of the joint distribution is

\[
M(t_1, t_2) = \int_0^\infty \int_x^\infty \exp(t_1 x + t_2 y - y) \, dy \, dx
\]

\[
= \frac{1}{(1 - t_1 - t_2)(1 - t_2)}
\]

for any \(t_2 < 1\) and \(t_1 + t_2 < 1\)

- For \(X\) we have \(M(t_1, 0) = \frac{1}{1 - t_1}\) and for \(Y\) \(M(0, t_2) = \frac{1}{(1 - t_2)^2}\)

- These correspond to \(f_1(x) = \exp(-x)\) and \(f_2(y) = y \exp(-y)\)