Multiple Random Variables

Lecture 3

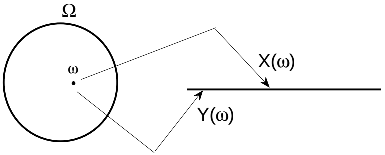
Sept. 9, 11 and 16, 2025

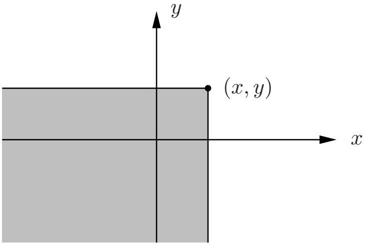
Outline

Two Random Variables

Two random variables X and Y, defined over the same probability space $(\Omega, \mathcal{F}, P)$, $X : \Omega \to \mathbb{R}$, $X : \Omega \to \mathbb{R}$ and $Y, : \Omega \to \mathbb{R}$, are specified by their joint cdf $(-\infty < x, y < \infty)$

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) = P(\{\omega \,|\, X(\omega) \leq x\} \cap \{\omega \,|\, Y(\omega) \leq y\})$$



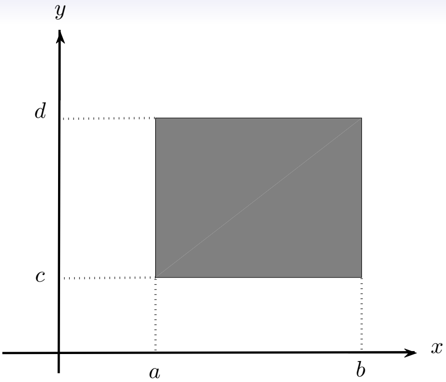


$F_{XY}(x, y)$ is the probability of the shaded region above.

Joint cdf

- Properties of the joint cdf
  - $F_{XY}(x, y) \geq 0, \ -\infty < x, y < \infty$
  - If $x_1 \leq x_2$ and $y_1 \leq y_2$, then $F_{XY}(x_1, y_1) \leq F_{XY}(x_2, y_2)$
  - $F_{XY}(x, -\infty) = 0 = F_{XY}(-\infty, y)$
  - $F_{XY}(\infty, y) = F_Y(y)$ and $F_{XY}(x, \infty) = F_X(x)$. $F_X(x)$ and $F_Y(y)$ are the marginal cdfs of X and Y, respectively.
  - $F_{XY}(\infty, \infty) = 1$
  - $F_{XY}(x, y) = \lim_{\epsilon \to 0, \, \delta \to 0} F_{XY}(x + \epsilon, y + \delta)$, $\epsilon, \delta > 0$: it is right continuous.
- X and Y are independent if for any x and y

$$F_{XY}(x, y) = F_X(x) F_Y(y)$$



$$P(a < X \le b,\ c < Y \le d)$$
$$= F_{XY}(b,d) - F_{XY}(a,d) - F_{XY}(b,c) + F_{XY}(a,c)$$

Joint, Marginal, and Conditional PMFs

- Let X and Y be discrete random variables on the same probability space
- They are completely specified by their joint pmf

$$P_{XY}(x,y) = P(X=x, Y=y), \quad x \in \mathcal{X},\ y \in \mathcal{Y}$$

By axioms of probability, $\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} P_{XY}(x,y) = 1$

- Example Consider the PMF $P_{XY}(x,y)$ given in the table

		x		
		0	1	2.5
y	-3	0	$\frac{1}{4}$	$\frac{1}{8}$
	-1	$\frac{1}{8}$	0	$\frac{1}{4}$
	2	$\frac{1}{8}$	$\frac{1}{8}$	0

- To find $P_X(x)$, the marginal pmf of X, we use total probability

$$P_X(x) = \sum_{y \in \mathcal{Y}} P_{XY}(x,y), \quad x \in \mathcal{X}$$

		x: 0	1	2.5	$P_Y(y)$
y	-3	0	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{3}{8}$
	-1	$\frac{1}{8}$	0	$\frac{1}{4}$	$\frac{3}{8}$
	2	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{4}$
	$P_X(x)$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	

- The conditional pmf of X given $Y = y$, is defined as

$$P_{X|Y}(x|y) = \frac{P_{XY}(x,y)}{P_Y(y)}, \quad P_Y(y) \neq 0,\ x \in \mathcal{X}$$

- Chain rule $P_{XY}(x,y) = P_Y(y)P_{X|Y}(x|y) = P_X(x)P_{Y|X}(y|x)$
- X and Y are independent if for any $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, $P_{XY}(x,y) = P_X(x)P_Y(y)$, which is equivalent to $P_{X|Y}(x|y) = P_X(x)$.

Joint, Marginal, and Conditional pdfs

- X and Y are jointly continuous random variables if their joint cdf is continuous in both x and y. In this case, we can define their joint pdf (probability density function), provided that it exists, as the function $f_{XY}(x, y)$ such that

$$F_{XY}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u, v)\, du\, dv, \quad x, y \in \mathbb{R}$$

- If $F_{XY}(x, y)$ is differentiable in x and y, then

$$\begin{aligned} f_{XY}(x, y) &= \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \\ &= \lim_{\Delta x, \Delta y \to 0} \frac{P(x < X \leq x + \Delta x, y < Y \leq y + \Delta y)}{\Delta x \Delta y} \end{aligned}$$

- Properties of $f_{XY}(x,y)$:
  - $f_{XY}(x,y) \geq 0$
  - $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1 = F_{XY}(\infty,\infty)$
- The marginal pdf of X, $f_X(x)$, is given by

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

$F_X(x) = F_{XY}(x,\infty) = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f_{XY}(u,y)\,dy\right)du$. Recall the Leibniz rule: $\frac{\partial}{\partial x}\int_{e(x)}^{h(x)} g(u,x)du = g(h(x),x)\frac{\partial h(x)}{\partial x} - g(e(x),x)\frac{\partial e(x)}{\partial x}$ $+ \int_{e(x)}^{h(x)}\frac{\partial g(u,x)}{\partial x}du$. So $f_X(x) = \frac{\partial}{\partial x}\int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f_{XY}(u,y)\,dy\right)du$ $= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$

- X and Y are independent if and only if $f_{XY}(x,y) = f_X(x)f_Y(y)$ for every x,y.

EXAMPLE

Let $(X, Y) \sim f_{XY}(x,y)$ where

$$f_{XY}(x,y) = \begin{cases} c & x \geq 0,\ y \geq 0,\ x+y \leq 1 \\ 0 & \mathit{otherwise}\,. \end{cases}$$

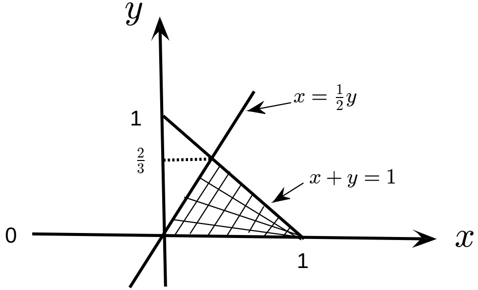
1) Find c
2) Find $f_Y(y)$
3) Are X and Y independent?
4) Find $P(X \geq 0.5Y)$.

1) $1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = \int_0^1 \int_0^{1-y} c\,dx\,dy = c\int_0^1 (1-y)dy = \frac{c}{2}$, hence, $c = 2$.
2) Use $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$:

$$f_Y(y) = \begin{cases} \int_0^{1-y} 2\,dx & 0 \leq y \leq 1 \\ 0 & \mathit{otherwise} \end{cases} = \begin{cases} 2(1-y) & 0 \leq y \leq 1 \\ 0 & \mathit{otherwise} \end{cases}$$

3) X and Y are independent iff $f_{XY}(x,y) = f_X(x) f_Y(y)$. Consider $(x,y) = (0,1)$. Then $f_{XY}(0,1) = c = 2$ but $f_X(0) f_Y(1) = 0$ (since $f_Y(1) = 0$). So X and Y are not independent.

4) To find the probability of the set $\{X \geq 0.5Y\}$. we first sketch it



From the figure we find that

$$
\begin{aligned}
P(X \geq 0.5Y) &= \int_{\{(x,y):x \geq 0.5y\}} f_{XY}(x,y)\,dx\,dy \\
&= \int_0^{\frac{2}{3}} \int_{\frac{y}{2}}^{1-y} 2\,dx\,dy = \int_0^{\frac{2}{3}} (2-3y)\,dy \\
&= \frac{2}{3}
\end{aligned}
$$

Conditional CDF and pdf

- Let X and Y be continuous random variables with joint pdf $f_{XY}(x,y)$. We wish to define $F_{Y|X}(y|X=x) = P(Y \leq y|X=x)$. We cannot define it as

$$\frac{P(Y \leq y,\, X = x)}{P(X = x)}$$

because both numerator and denominator are equal to zero.

- Instead, we define conditional probability for continuous random variables as a limit (if $f_X(x) \neq 0$)

$$\begin{aligned}
F_{Y|X}(y|X=x) &= \lim_{\Delta x \to 0} P(Y \leq y | x < X \leq x + \Delta x) \\
&= \lim_{\Delta x \to 0} \frac{P(Y \leq y,\ x < X \leq x + \Delta x)}{P(x < X \leq x + \Delta x)} \\
&= \lim_{\Delta x \to 0} \frac{\int_{-\infty}^{y} f_{XY}(x,u)\, du\, \Delta x}{f_X(x) \Delta x} = \frac{\int_{-\infty}^{y} f_{XY}(x,u)\, du}{f_X(x)}
\end{aligned}$$

- We then define the conditional pdf in the usual way as

$$f_{Y|X}(y|x) = \frac{\partial F_{Y|X}(y|X=x)}{\partial y} = \frac{f_{XY}(x,y)}{f_X(x)} \text{ if } f_X(x) \neq 0$$

We will write the above as $Y \mid \{X = x\} \sim f_{Y|X}(y|x)$

Example: Let $(X, Y) \sim f_{XY}(x, y)$ where

$$f_{XY}(x,y) = \begin{cases} 2 & x \geq 0,\ y \geq 0,\ x+y \leq 1 \\ 0 & \textit{otherwise}\,. \end{cases}$$

Find $f_{X|Y}(x|y)$
Earlier we derived the solution

$$f_Y(y) = \begin{cases} 2(1-y) & 0 \leq y \leq 1 \\ 0 & \textit{otherwise} \end{cases}$$

$$\Rightarrow f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{1-y} & 0 \leq y \leq 1,\ 0 \leq x \leq 1-y \\ 0 & \textit{otherwise} \end{cases}$$

That is, we have a uniform distribution: $X \mid \{Y = y\} \sim U[0, 1-y]$

Expectation

- Let $(X, Y) \sim f_{XY}(x, y)$ and let $g(x, y)$ be a function of x and y. The expectation (or expected value or mean) of $g(X, Y)$ is given by

$$E\{g(X, Y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y)\, dx\, dy$$

  The function $g(X, Y)$ may be X, Y, X^2, $X + Y$, etc.
- If X and Y are discrete, $(X, Y) \sim P_{XY}(x, y)$, then

$$E\{g(X, Y)\} = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} g(x, y) P_{XY}(x, y)$$

- The correlation of X and Y is defined as $E\{XY\}$.
- The covariance of X and Y is defined as

$$\begin{aligned} \mathrm{Cov}(X, Y) =& E\{(X - E\{X\})(Y - E\{Y\})\} \\ =& E\{XY - XE\{Y\} - YE\{X\} + E\{X\}E\{Y\}\} \\ =& E\{XY\} - E\{X\}E\{Y\} \end{aligned}$$

- Note that $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

Example: Let $(X, Y) \sim f_{XY}(x, y)$ where

$$f_{XY}(x,y) = \begin{cases} 2 & x \geq 0,\ y \geq 0,\ x+y \leq 1 \\ 0 & otherwise\,. \end{cases}$$

Find $E\{X\}$, $\mathrm{Var}(X)$ and $\mathrm{Cov}(X, Y)$

$$\begin{aligned} E\{X\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{XY}(x,y)\, dx\, dy \\ &= \int_0^1 2x \left(\int_0^{1-x} dy\right)\, dx = 2\int_0^1 (1-x)x\, dx = 2(\frac{1}{2} - \frac{1}{3}) = \frac{1}{3} \end{aligned}$$

Since $\mathrm{Var}(X) = E\{X^2\} - (E\{X\})^2$, we need to find the second moment

$$E\{X^2\} = 2\int_0^1 (1-x)x^2\, dx = 2(\frac{1}{3} - \frac{1}{4}) = \frac{1}{6}\,.$$

Hence, $\mathrm{Var}(X) = \frac{1}{6} - \frac{1}{3^2} = \frac{1}{18}$

By symmetry $E\{Y\} = E\{X\} = \frac{1}{3}$. Thus the covariance of X and Y is

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x, y)\, dx\, dy - E\{X\}E\{Y\} \\
&= \int_0^1 2x \left(\int_0^{1-x} ydy\right)\, dx - \frac{1}{3^2} = \int_0^1 x(1-x)^2\, dx - \frac{1}{9} \\
&= \frac{1}{12} - \frac{1}{9} = -\frac{1}{36}
\end{aligned}$$

Bounding Probability Using Expectation

- In many cases we do not know the distribution of an r.v. X but want to find the probability of an event such as $\{X > a\}$ or $\{|X - E\{X\}| > a\}$
- The Markov and Chebyshev inequalities give upper bounds on the probabilities of such events in terms of the mean and variance of the random variable
- Example: Let $X \geq 0$ represent the age of a person in the Atlanta Area. If we know that $E\{X\} = 35$ years, what fraction of the population is ≥ 70 years old?
  Clearly we cannot answer this question knowing only the mean, but we can say that $P(X \geq 70\} \leq 0.5$, since otherwise the mean would be larger than 35.
- This is an application of the Markov inequality

Markov Inequality

Markov inequality: For any r.v. $X \geq 0$ with finite mean $E\{X\}$ and any $a > 0$,

$$P(X \geq aE\{X\}) \leq \frac{1}{a}$$

Proof:

$$\begin{aligned}
E\{X\} &= \int_0^\infty x f_X(x)\, dx \quad \text{since } X \geq 0 \\
&= \underbrace{\int_0^a x f_X(x)\, dx}_{\geq 0} + \int_a^\infty x f_X(x)\, dx \\
&\geq \int_a^\infty x f_X(x)\, dx \geq \int_a^\infty a f_X(x)\, dx \\
&= aP(X \geq a)
\end{aligned}$$

$$\Rightarrow P(X \geq a) \leq \frac{E\{X\}}{a}$$

- The Markov inequality can be very loose. If $X \sim \exp(1)$, then

$$P(X \geq 10) = e^{-10} \approx 4.54 \times 10^{-5}$$

The Markov inequality gives

$$P(X \geq 10) \leq \frac{E\{X\}}{10} = \frac{1}{10},$$

which is very pessimistic.

- But it is the tightest possible bound on $P(X \geq aE\{X\})$ when we are given only the mean of X.
To show this, note that the inequality is tight for the following r.v.:

$$X = \begin{cases} aE\{X\} & \text{with probability } 1/a \\ 0 & \text{with probability } 1 - 1/a \end{cases}$$

Chebyshev Inequality

- Let X be a device parameter in an integrated circuit (IC) with known mean and variance. The IC is out-of-spec if X is more than, say, $3\sigma_X$ away from its mean. We wish to find the fraction of out-of-spec ICs, namely, $P(|X - E\{X\}| > 3\sigma_X)$.
  The Chebyshev inequality gives us an upper bound on this fraction in terms the mean and variance of X
- Chebyshev inequality: For any r.v. X with finite mean $E\{X\}$ and variance $\mathrm{Var}(X) = \sigma_X^2$,

$$P(|X - E\{X\}| > a\sigma_X) \leq \frac{1}{a^2}$$

  Proof: We use the Markov inequality. Define $Y = (X - E\{X\})^2 \geq 0$. Since $E\{Y\} = \sigma_X^2$, the Markov inequality gives

$$P(Y > a^2\sigma_X^2) \leq \frac{1}{a^2}\,.$$

  But $\{Y > a^2\sigma_X^2\}$ is equivalent to $\{|X - E\{X\}| > a\sigma_X\}$. Hence $P(|X - E\{X\}| > a\sigma_X) \leq \frac{1}{a^2}$.
- An alternative form (Gubner p. 89) is $P(|Y| > a) \leq E\{Y^2\}/a^2$ which follows from Markov: $P(|Y| > a) = P(|Y|^2 > a^2)$

- The Chebyshev inequality can be very loose. If $X \sim \mathcal{N}(0,1)$, then

$$P(|X| \geq 3) \approx 2Q(3) = 2 \times 10^{-3}$$

The Chebyshev inequality gives

$$P(|X| \geq 3) \leq \frac{\sigma_X^2}{3^2} = \frac{1}{9},$$

which is very pessimistic compared to 2×10^{-3}.

- But it is the tightest possible bound on $P(|X - E\{X\}| > a\sigma_X)$ when we are given only the mean and variance of X.
To show this, note that the equality is achieved for the following r.v.:

$$X = \begin{cases} E\{X\} + a\sigma_X & \text{with probability } \frac{1}{2a^2} \\ E\{X\} - a\sigma_X & \text{with probability } \frac{1}{2a^2} \\ E\{X\} & \text{with probability } 1 - \frac{1}{a^2} \end{cases}$$

Chernoff Bound

Chernoff Bound: For any r.v. X with finite mean $E\{X\}$ and any $a > 0$,

$$P(X \geq a) \leq \min_{s \geq 0} \left[e^{-as} E\{e^{sX}\} \right]$$

Proof:

$$\begin{aligned} P(X \geq a) =& P(sX \geq sa) \quad \text{for any } s \geq 0 \\ =& P\left(\underbrace{e^{sX}}_{Y} \geq \underbrace{e^{sa}}_{b} \right) \overset{Markov}{\leq} \frac{E\{Y\}}{b} = e^{-as} E\{e^{sX}\} \end{aligned}$$

The right-side is true for any $s \geq 0$; minimize it w.r.t. $s \geq 0$ to obtain the tightest bound. $\square$

$M_X(s) = E\{e^{sX}\} = \int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx$ is called the Moment Generating Function (mgf) of X. (Gubner p. 156). Here we allow s to be complex-valued, and notice that $M_X(-s)$ is the two-sided Laplace Transform of the pdf of X. Invoking the properties of the Laplace transform, one has one-to-one relationship between the pdf of X and mgf of X!

Example (p. 166): Consider $X \sim \exp(1)$. Find $P(X \geq 7)$, and Markov, Chebyshev and Chernoff bounds.

Solution: Note that $E\{X\} = \lambda^{-1} = 1$, $\text{var}(X) = \sigma_X^2 = \lambda^{-2} = 1$ and $E\{X^2\} = 2\lambda^{-2} = 2$

- Exact: $P(X \geq 7) = \int_7^\infty e^{-x}\, dx = e^{-7} = 0.00091$
- Markov: $P(X \geq 7) \leq \frac{E\{X\}}{7} = \frac{1}{7} = 0.143$
- Chebyshev: $P(X \geq 7) = P(|X| \geq 7) \leq \frac{E\{X^2\}}{7^2} = \frac{2}{49} = 0.041$
- Chernoff:

$$E\{e^{sX}\} = \int_0^\infty e^{sx} e^{-x}\, dx = \frac{e^{(s-1)x}}{s-1}\Big|_{x=0}^{\infty} = \frac{1}{1-s} \text{ if } 1 - s > 0$$

Now $P(X \geq 7) \leq \min_{0 \leq s < 1} g(s)$ where $g(s) = e^{-7s}/(1-s)$. We have

$$\frac{dg(s)}{ds} = \frac{(1-s)(-7e^{-7s}) - e^{-7s}(-1)}{(1-s)^2} = 0$$

$$\Rightarrow -6 + 7s = 0 \Rightarrow s = \frac{6}{7}$$

Hence, $P(X \geq 7) \leq g(6/7) = 7e^{-6} = 0.017$