Random Variables and Expectation (III)

Josep Díaz    Maria J. Serna    Conrado Martínez
U. Politècnica de Catalunya

RA-MIRI 2022–2023
Recall $f : \mathbb{R} \to \mathbb{R}$ is **convex** if, for all $x_1, x_2 \in \mathbb{R}$ and for all $t \in [0, 1]$, we have

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

If $f$ is twice differentiable, a necessary and sufficient condition for $f$ to be convex is that $f''(x) \geq 0$ for $x \in \mathbb{R}$.

**Lemma**

*If $f$ is convex then $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$.***
Jensen’s inequality

Proof

Let $\mu = \mathbb{E}[X]$ ($\mu \in \mathbb{R}$). Using Taylor to expand $f$ at $X = \mu$,

$$f(X) = f(\mu) + f'(\mu)(X - \mu) + \frac{f''(\mu)(X - \mu)^2}{2} + \cdots$$

$$\geq f(\mu) + f'(\mu)(X - \mu)$$

$$\mathbb{E}[f(X)] \geq \mathbb{E}[f(\mu) + f'(\mu)(X - \mu)]$$

$$= \mathbb{E}[f(\mu)] + f'(\mu)(\mathbb{E}[X] - \mu) = f(\mu)$$

i.e., $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$. \qed
Expectation of combinations of r.v.

Consider the following experiment:
\( X = \text{Uniform}\{1, 2\} \) and \( Y = \text{Uniform}\{1, X + 1\} \)
Thus \( Y \) depends on \( X \).
What is the expectation of the r.v. \( XY \)?

\[ \Omega = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3)\} \]

\[
E[XY] = \sum_{\omega \in \Omega} X(\omega)Y(\omega)P[\omega]
\]

We have

\[
P[(1, 1)] = P[(1, 2)] = \frac{1}{4} ;
\]
\[
P[(2, 1)] = P[(2, 2)] = P[(2, 3)] = \frac{1}{6}.
\]

\[
E[XY] = \frac{1}{4} \cdot 1 \cdot 1 + \frac{1}{4} \cdot 1 \cdot 2 + \frac{1}{6} \cdot 2 \cdot 1 + \frac{1}{6} \cdot 2 \cdot 2 + \frac{1}{6} \cdot 2 \cdot 3 = \frac{11}{4}.
\]
Expectation of combinations of r.v.

Consider the following experiment:

\( X = \text{Uniform}\{1, 2\} \) and \( Y = \text{Uniform}\{1, X + 1\} \)

Thus \( Y \) depends on \( X \).

What is the expectation of the r.v. \( XY \)?

\( \Omega = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3)\} \)

\[
E[XY] = \sum_{\omega \in \Omega} X(\omega)Y(\omega) \mathbb{P}[\omega]
\]

We have

\[
\mathbb{P}[(1, 1)] = \mathbb{P}[(1, 2)] = 1/4;
\]

\[
\mathbb{P}[(2, 1)] = \mathbb{P}[(2, 2)] = \mathbb{P}[(2, 3)] = 1/6.
\]

\[
E[XY] = \frac{1}{4} \cdot 1 \cdot 1 + \frac{1}{4} \cdot 1 \cdot 2 + \frac{1}{6} \cdot 2 \cdot 1 + \frac{1}{6} \cdot 2 \cdot 2 + \frac{1}{6} \cdot 2 \cdot 3 = \frac{11}{4}.
\]
Expectation of combinations of r.v.

Consider the following experiment:
\( X = \text{Uniform}(\{1, 2\}) \) and \( Y = \text{Uniform}(\{1, X + 1\}) \)
Thus \( Y \) depends on \( X \).
What is the expectation of the r.v. \( XY \)?

\[ \Omega = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3)\} \]

\[
\mathbb{E}[XY] = \sum_{\omega \in \Omega} X(\omega)Y(\omega) \mathbb{P}[\omega]
\]

We have
\[
\mathbb{P}[(1, 1)] = \mathbb{P}[(1, 2)] = \frac{1}{4}; \\
\mathbb{P}[(2, 1)] = \mathbb{P}[(2, 2)] = \mathbb{P}[(2, 3)] = \frac{1}{6}.
\]

\[
\mathbb{E}[XY] = \frac{1}{4} \cdot 1 \cdot 1 + \frac{1}{4} \cdot 1 \cdot 2 + \frac{1}{6} \cdot 2 \cdot 1 + \frac{1}{6} \cdot 2 \cdot 2 + \frac{1}{6} \cdot 2 \cdot 3 = \frac{11}{4}.
\]
We have, \( \mathbb{P}[X = 1] = 1/2 \); \( \mathbb{P}[X = 2] = 1/2 \) and

\[
\begin{align*}
\mathbb{P}[Y = 1] &= \mathbb{P}[Y = 1|X = 1] \mathbb{P}[X = 1] + \mathbb{P}[Y = 1|X = 2] \mathbb{P}[X = 2] = 1/4 + 1/6 = 5/12; \\
\mathbb{P}[Y = 2] &= \mathbb{P}[Y = 2|X = 1] \mathbb{P}[X = 1] + \mathbb{P}[Y = 2|X = 2] \mathbb{P}[X = 2] = 1/4 + 1/6 = 5/12; \\
\end{align*}
\]

Then \( \mathbb{E}[X] = 3/2 \) and \( \mathbb{E}[Y] = 7/4 \) so \( \mathbb{E}[X] \mathbb{E}[Y] = 21/8 \). Therefore,

\[
\mathbb{E}[XY] \neq \mathbb{E}[X] \mathbb{E}[Y].
\]
Joint Probability Mass Function

The joint PMF of r.v. \(X, Y\) is the function \(p_{XY} : \mathbb{R}^2 \rightarrow \mathbb{R}\) defined by \(p_{XY}(x, y) = \mathbb{P}[X = x \land Y = y]\).

With the joint PMF of r.v. \(X, Y\) you can compute the expectation of any function \(f(X, Y)\):

\[
\mathbb{E}[f(X, Y)] = \sum_{x, y} f(x, y) \cdot p_{XY}(x, y).
\]

Compute \(\mathbb{E}\left[\frac{X}{Y}\right]\) for the previous r.v. \(X, Y\)

\[
\mathbb{E}\left[\frac{X}{Y}\right] = p_{XY}(1, 1) \frac{1}{1} + p_{XY}(1, 2) \frac{1}{2}
\]
\[
+ p_{XY}(2, 1) \frac{2}{1} + p_{XY}(2, 2) \frac{2}{2} + p_{XY}(2, 3) \frac{2}{3}
\]
\[
= \frac{1}{4} \cdot (1 + 1/2) + \frac{1}{3} \cdot (2 + 1 + 2/3) = \frac{3}{8} + \frac{11}{3} = \frac{97}{24} = 4 \frac{1}{24}
\]
Independent r.v.: Main result

**Theorem**

If $X$ and $Y$ are independent r.v. then $\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]$.

**Proof**

\[
\mathbb{E}[X \cdot Y] = \sum_{x,y} p_{XY}(x, y) \cdot x \cdot y \\
= \sum_{x,y} p_X(x) \cdot p_Y(y) \cdot x \cdot y \quad \text{(by independence)} \\
= \sum_{x,y} x \cdot p_X(x) \cdot y \cdot p_Y(y) \\
= \left( \sum_x x \cdot p_X(x) \right) \cdot \left( \sum_y y \cdot p_Y(y) \right) \\
= \mathbb{E}[X] \cdot \mathbb{E}[Y]
\]
The Poisson approximation to the Binomial

For $X \sim \text{Bin}(n, p)$, for large $n$, computing the PMF $\mathbb{P}[X = x]$ could be quite nasty.

It turns out that for large $n$ and small $p$, $\text{Bin}(n, p)$ can be easily approximated by the PMF of a simpler Poisson random variable.

A discrete r.v. $X$ is Poisson with parameter $\lambda$ ($X \sim \text{Poisson}(\lambda)$), if it has PMF $\mathbb{P}[X = i] = \frac{\lambda^i e^{-\lambda}}{i!}$, for $i \in \{0, 1, 2, 3, \ldots\}$

If $X \sim \text{Poisson}(\lambda)$ then $\mathbb{E}[X] = \lambda$.

This is the reason that sometimes $\lambda$ is denoted $\mu$.

Proof

$$
\mathbb{E}[X] = \sum_{i=1}^{\infty} i \frac{\lambda^i e^{-\lambda}}{i!} = e^{-\lambda}\lambda \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = e^{-\lambda}\lambda e^\lambda = \lambda
$$

Taylor for $e^\lambda$
The Poisson approximation to the Binomial

**Theorem**

If $X \in Bin(n, p)$, with $\mu = np$, then as $n \to \infty$, for each fixed $i \in \{0, 1, 2, 3, \ldots\}$,

$$
\mathbb{P}[X = i] \sim \frac{\mu^i e^{-\mu}}{i!}.
$$

**Proof**

As $\mu = np$,

$$
\mathbb{P}[X = i] = \binom{n}{i} \left(\frac{\mu}{n}\right)^i \left(1 - \frac{\mu}{n}\right)^{n-i}
$$

$$
= \frac{n(n-1) \cdots (n-i+1) \mu^i}{i!} \frac{(1 - \frac{\mu}{n})^n}{n^i} \left(1 - \frac{\mu}{n}\right)^{-i}
$$

$$
= \frac{\mu^i}{i!} \left(1 - \frac{\mu}{n}\right)^n \frac{n(n-1) \cdots (n-i+1)}{n^i} \left(1 - \frac{\mu}{n}\right)^{-i}
$$

$$
\sim \frac{\mu^i}{i!} e^{-\mu} \text{ as } n \to \infty.
$$
Example

The population of Catalonia is around 7 million people. Assume that the probability that a person is killed by lightning in a year is \( p = \frac{1}{5 \times 10^8} \).

a) Let's compute the exact probability that 3 or more people will be killed by lightning next year in Catalonia.
Let \( X \) be a r.v. counting the number of people that will be killed in Cat. next year by a lightning.
We want to compute
\[
P[X \geq 3] = 1 - P[X = 0] - P[X = 1] - P[X = 2],
\]
where \( X \sim \text{Bin}(7 \times 10^6, \frac{1}{5 \times 10^8}) \).
Then,
\[
P[X \geq 3] = 1 - (1 - p)^n - np(1 - p)^{n-1} - \binom{n}{2} p^2 (1 - p)^{n-2} = 1.65422 \times 10^{-7}
\]
Example

b) Use Poisson approximation to approximate $P[X \geq 3]$. 
$\lambda = np = 7/500$ so 

$$P[X \geq 3] \sim 1 - e^{\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2}{2} e^{-\lambda} = 1.52558 \times 10^{-7}$$

c) Approximate the probability that 2 or more people will be killed by lightning the first 6 months of the year 
Notice we are considering $\lambda$ as a rate. Then we have now 
$\lambda = (7/500)/2$ 

$$P[X \geq 2 \text{ during 6 months}] \sim 1 - e^{\lambda} - \lambda e^{-\lambda} = 5.79086 \times 10^{-7}$$

d) Approximate the probability that in 3 of the next 10 years exactly 3 people will be killed 
We have $\lambda = 7/500$, then the probability that in any particular year 3 people are killed is 

$$= \frac{e^{-\lambda} \lambda^3}{3!}.$$ 
Let $Y$ be a r.v. counting the number of years with exactly 3 kills. 
Assuming independence between years, $Y \sim \text{Bin}(10, \frac{e^{-\lambda} \lambda^3}{3!})$, 
therefore the answer is 

$$\binom{10}{3} \left( \frac{e^{-\lambda} \lambda^3}{3!} \right)^3 \left( 1 - \frac{e^{-\lambda} \lambda^3}{3!} \right)^7 \approx 1.1 \cdot 10^{-17}$$