Math 170E

Lecture Notes Section 5.3 [*,†]

Independent random variables

Instructor: Swee Hong Chan

NOTE: Materials that appear in the textbook but do not appear in the lecture notes might still be tested. Please send me an email if you find typos.

*Version date: Wednesday 3rd March, 2021, 15:00.

†This notes is based on March Boedihardjo's and Jamie Haddock's notes from the past quarters, and I would like to thank them for their generosity. "*Nanos gigantum humeris insidentes* (I am but a dwarf standing on the shoulders of giants)".

1 Functions of independent random variables

We want to solve the following problem:

- Input: **Independent** random variables $X_1, X_2, \ldots, X_n$ and functions $u_1(\cdot), u_2(\cdot), \ldots, u_n(\cdot)$.

- Output: A new random variable

$$Y = u_1(X_1) + u_2(X_2) + \ldots + u_n(X_n),$$

or

$$Y = u_1(X_1)\,u_2(X_2) \ldots u_n(X_n).$$

2 Independent random variables: Example

The time until defaults of a group of credit card holders follows an exponential distribution with mean 2 years.

The credit card company will go bankrupt if every one of this group defaults within one year.

Determine the probability that this will happen, if there are 100 people in this group.

3 Independent random variables: Answer

Let $X_1, X_2, \ldots, X_{100}$ be the time until default of the 100 members.

We assume that these RVs are **independent**, since the members of this group are strangers to each other.

Recall that, for **independent** RVs, the joint cdf is the **product** of its marginal cdfs. Thus

$$
\begin{aligned}
P(X_1 \le 1, \ldots, X_{100} \le 1) &= P(X_1 \le 1) \ldots P(X_{100} \le 1) \\
= \left(1 - e^{-\frac{1}{2}}\right) \ldots \left(1 - e^{-\frac{1}{2}}\right) &= (1 - e^{-\frac{1}{2}})^{100} \approx 3.1 \times 10^{-41}.
\end{aligned}
$$

Note: Since on average it took **2 years** to default, it is extremely unlikely that everyone defaults within **1 year**.

4 Product formula: Theorem

Theorem 1. *Let $X_1, X_2, \ldots, X_n$ be* ***independent*** *RVs, let $u_1(\cdot), u_2(\cdot), \ldots, u_n(\cdot)$ be functions, and let*

$$Y = u_1(X_1)\, u_2(X_2)\ \ldots\ u_n(X_n).$$

Then

$$E[Y] \quad = \quad E[u_1(X_1)]\ E[u_2(X_2)]\ \ldots\ E[u_n(X_n)].$$

5 Product formula: Example

Let X_1 be binomial RV with $n_1 = 4, p_1 = \frac{1}{2}$.

Let X_2 be binomial RV with $n_2 = 6, p_2 = \frac{1}{3}$.

Let X_3 be binomial RV with $n_3 = 12, p_3 = \frac{1}{6}$.

Suppose that X_1, X_2, X_3 are independent.

Compute $E[X_1 X_2 X_3]$.

Answer: We have

$$
\begin{aligned}
E[X_1\, X_2\, X_3] &= E[X_1]\, E[X_2]\, E[X_3] \\
&= (n_1 p_1)\ (n_2 p_2)\ (n_3 p_3) \ = \left(4 \times \frac{1}{2}\right) \left(6 \times \frac{1}{3}\right) \left(12 \times \frac{1}{6}\right) \\
&= 8.
\end{aligned}
$$

6 Addition formula: Theorem

Theorem 2. *Let $X_1, X_2, \ldots, X_n$ be (not necessarily independent) RVs, let $a_1, \ldots, a_n$ be real numbers, and let*

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n.$$

Then

$$E[Y] = a_1 E[X_1] + a_2 E[X_2] + \ldots + a_n E[X_n].$$

*Suppose now that $X_1, X_2, \ldots, X_n$ are **independent**. Then*

$$\sigma^2(Y) = a_1^2 \sigma^2(X_1) + a_2^2 \sigma^2(X_2) + \ldots + a_n^2 \sigma^2(X_n).$$

7 Addition formula: Example

Let X_1 be binomial RV with $n_1 = 4, p_1 = \frac{1}{2}$.

Let X_2 be binomial RV with $n_2 = 6, p_2 = \frac{1}{3}$.

Let X_3 be binomial RV with $n_3 = 12, p_3 = \frac{1}{6}$.

Suppose that X_1, X_2, X_3 are independent.

Find the mean and variance of $Y = X_1 + 2X_2 + 3X_3$.

8 Addition formula: Answer

We have

$$
\begin{aligned}
E[X_1 + 2X_2 + 3X_3] \quad &= \quad E[X_1] + 2E[X_2] + 3E[X_3] \\
= \quad & n_1 p_1 \ + \ 2n_2 p_2 \ + \ 3n_3 p_3 \\
= \quad & \left(4 \times \frac{1}{2}\right) \ + \ 2\left(6 \times \frac{1}{3}\right) \ + \ 3\left(12 \times \frac{1}{6}\right) \\
= \quad & 12,
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma^2(Y) \quad &= \quad \sigma^2(X_1) + 4\sigma^2(X_2) + 9\sigma^2(X_3) \\
= \quad & n_1 p_1 q_1 \ + \ 4n_2 p_2 q_2 \ + \ 9n_3 p_3 q_3 \\
= \quad & \left(4 \times \frac{1}{2} \times \frac{1}{2}\right) \ + \ 4\left(6 \times \frac{1}{3} \times \frac{2}{3}\right) \ + \ 9\left(12 \times \frac{1}{6} \times \frac{5}{6}\right) \\
= \quad & 1 + \frac{16}{3} + 15 \quad = \quad 19\frac{1}{3}.
\end{aligned}
$$

9 (Super weak) law of large numbers: Theorem

Theorem 3. *Let* $X_1, X_2, \ldots, X_n$ *be independent random variables with mean* μ *and variance* σ^2*. Let*

$$\overline{\mathrm{X}} := \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Then

$$E[\overline{\mathrm{X}}] = \mu; \qquad Var(\overline{\mathrm{X}}) = \frac{\sigma^2}{n}.$$

That means, as $n \to \infty$, the **mean** remains unchanged while the **variance** decreased to 0.

Thus we see **almost no randomness** when n is large.

10 Law of large numbers: Intuition

The result of one dice throw is unforeseeable , the result of one million dice throws is predictable.

The law of large number says that, the **average** of the results obtained from a **large number** of trials should be close to the **expected value**.

This is something that we would expect based on intuition, but the law of large number allows us to express this intuition using precise mathematical formula.

11 Law of large numbers: Applications

• Every game in casinos is designed so that the house has a slight edge (i.e., their expected winning is positive). Law of large number thus says, as long as casinos have many customers (thus why they advertise really hard) and the probabilities are all independent (i.e., no cheating), the casino will make a profit in the long run.

• The (fictitious) field of *psychohistory* in the *Foundation* series of Isaac Asimov could predict the general flow of future events of a civilization, provided that the population whose behavior was modeled is sufficiently large.