

Probability Theory and Statistics

Lecture 8

Bence Csonka

Budapest University of Technology and Economics
csonkab@edu.bme.hu

October 13

Convolution in the discrete case

If X and Y are independent random variables, what is the distribution of $X + Y$?

Theorem

Let X and Y be independent, *discrete* random variables taking values in the nonnegative integers. Then

$$\mathbb{P}(X + Y = k) = \sum_{i=0}^k \mathbb{P}(X = i) \mathbb{P}(Y = k - i)$$

for every $k \in \{0, 1, 2, \dots\}$.

(Proof.)

(**Convolution**: the distribution of the sum of independent random variables.)

(For general discrete X, Y the formula is analogous.)

Convolution of Poisson

Let X and Y be independent random variables with $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$. Then $X + Y \sim \text{Pois}(\lambda + \mu)$, i.e., for any $k \in \{0, 1, 2, \dots\}$,

$$\mathbb{P}(X + Y = k) = \frac{(\lambda + \mu)^k}{k!} e^{-(\lambda + \mu)}.$$

(Proof: see the lecture.)

Definition

Let $X : \Omega \rightarrow \mathbb{R}$ be a continuous random variable. The distribution function of X is defined by

$$F_X(x) = \mathbb{P}(X < x).$$

The density function of X is the function which satisfies

$$F_X(x) = \int_{-\infty}^x f_X(t) \, dt.$$

Remainder: the Uniform Distribution

Definition

The density function of a random variable X uniformly distributed on the interval (a, b) is

$$f_X(x) = \frac{1}{b-a},$$

if $a < x < b$, and $f_X(x) = 0$ otherwise. By integration, we obtain

$$F_X(x) = \begin{cases} 0, & \text{if } x \leq a, \\ \frac{x-a}{b-a}, & \text{if } a < x < b, \\ 1, & \text{if } x \geq b. \end{cases}$$

Notation: $X \sim U(a; b)$. For the $U(a; b)$ distribution the density function is constant (on (a, b) , and zero outside), while the distribution function is linear (on (a, b) , and constant outside).

Remainder: Exponential Distribution

Definition

A random variable Z has an *exponential distribution* with parameter $\lambda > 0$ if

$$f_Z(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Equivalently,

$$F_Z(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1 - e^{-\lambda x}, & \text{if } x > 0. \end{cases}$$

Notation: $X \sim \text{Exp}(\lambda)$.

Occurrence: waiting times ($Z \geq 0$ with probability 1), e.g. the remaining lifetime of a light bulb until it burns out.

Example for Exponential Distribution

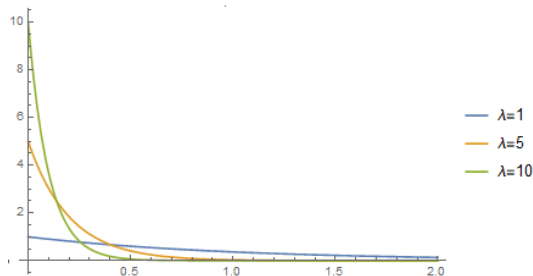
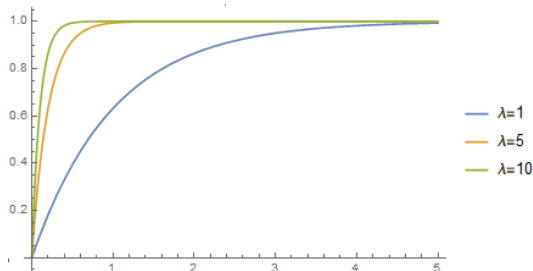
Let X be an exponentially distributed random variable such that $\mathbb{P}(X > 3) = e^{-6}$.

Example for Exponential Distribution

Let X be an exponentially distributed random variable such that $\mathbb{P}(X > 3) = e^{-6}$.

- a) What is the parameter λ of the distribution of X ?
- b) Compute $\mathbb{P}(X < 2)$.

Exponential Distribution: CDF and PDF



Exponential Distribution: Memorylessness

Reminder: the geometric distribution is memoryless: if $X \sim \text{Geo}(p)$, then

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\}. \quad (1)$$

Note: if at least one of s, t is not an integer, then (1) need not hold. For instance, let $t = 3/4$, $s = 1/2$, then

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > 5/4 \mid X > 3/4) = \mathbb{P}(X \geq 2) = 1 - p,$$

but

$$\mathbb{P}(X > s) = \mathbb{P}(X > 1/2) = 1.$$

However, if $X \sim \text{Exp}(\lambda)$ with $\lambda > 0$, then (1) holds for all $s, t \in [0, \infty)$.

Exponential Distribution: Memorylessness

Reminder: the geometric distribution is memoryless: if $X \sim \text{Geo}(p)$, then

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\}. \quad (1)$$

Note: if at least one of s, t is not an integer, then (1) need not hold. For instance, let $t = 3/4$, $s = 1/2$, then

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > 5/4 \mid X > 3/4) = \mathbb{P}(X \geq 2) = 1 - p,$$

but

$$\mathbb{P}(X > s) = \mathbb{P}(X > 1/2) = 1.$$

However, if $X \sim \text{Exp}(\lambda)$ with $\lambda > 0$, then (1) holds for all $s, t \in [0, \infty)$.

Modeling question: for which types of waiting times does (1) hold?

E.g. for light bulbs, measurements suggest it is *approximately* true. For human lifetimes, it is not true!

Memorylessness: Exponential vs. Geometric

Reminder:

Theorem

If a non-constant random variable X satisfies

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\} \quad (2)$$

and $\text{Ran}(X) = \{1, 2, \dots\}$, then $X \sim \text{Geo}(p)$ with $p = \mathbb{P}(X = 1)$.

Now:

Theorem

If a non-constant random variable X satisfies

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > s), \quad \forall s, t \geq 0 \quad (3)$$

and $\text{Ran}(X) = [0, \infty)$, then the distribution of X is exponential.

Proof of Memorylessness

Theorem

If a non-constant random variable X satisfies

$$\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > s), \quad \forall s, t \geq 0 \quad (4)$$

and $\text{Ran}(X) = [0, \infty)$, then the distribution of X is exponential.

(Proof)

In the proof we use the following lemma (stated without proof):

Lemma

If $g: [0, 1) \rightarrow \mathbb{R}$ is a non-increasing function such that $g(t + s) = g(t) + g(s)$ for all $s, t \in [0, \infty)$, then $g(t) = -\lambda t$, where $\lambda = -g(1) \geq 0$.

Link Between Exponential and Geometric Distributions

Due to memorylessness, the exponential distribution is the continuous analogue of the geometric distribution. By the next statement, the geometric distribution can actually be expressed via the exponential.

Theorem

If X is exponentially distributed with parameter λ , then $\lceil X \rceil$ is geometrically distributed with parameter $1 - e^{-\lambda}$.

(Proof.)

Link Between Exponential and Geometric Distributions

Due to memorylessness, the exponential distribution is the continuous analogue of the geometric distribution. By the next statement, the geometric distribution can actually be expressed via the exponential.

Theorem

If X is exponentially distributed with parameter λ , then $\lceil X \rceil$ is geometrically distributed with parameter $1 - e^{-\lambda}$.

(Proof.)

Note that there is also a connection between the exponential and the **Poisson** distributions: if $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Pois}(\lambda)$, then

$$\mathbb{P}(X \geq 1) = e^{-\lambda} = \mathbb{P}(Y = 0).$$

Not Every Random Variable is Discrete or Continuous...

Example: define X as follows. Toss a fair coin. If it lands heads, sample a number uniformly on $(0, 1)$ using your favorite random number generator and take this value as X ; if tails, roll a fair die and take the outcome as X . Then

- X is not continuous, because if it were, we would have $\mathbb{P}(X = 6) = 0$, but actually $\mathbb{P}(X = 6) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$ (multiplication rule).
- X is not discrete either, since on $(0, 1)$ it can take every value, hence its range is uncountable.

Conditional on heads, X behaves like a continuous r.v.; conditional on tails, like a discrete one.

Expectation in the Continuous Case

Reminder:

Definition

If X is a *nonnegative* discrete random variable with

$$\sum_{k \in \text{Ran}(X)} |k| \mathbb{P}(X = k) < \infty,$$

then

$$\mathbb{E}(X) = \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k).$$

The continuous analogue: instead of summing the weight function times k , we integrate the density times x —this is also a “probability-weighted average”.

Expectation in the Continuous Case

The continuous analogue: instead of summing the weight function times k , we integrate the density times x —this is also a “probability-weighted average”.

Definition

Let X be a continuous random variable such that

$$\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty.$$

Then

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Expected Values of Notable Distributions

Theorem

If $X \sim \text{Exp}(\lambda)$ for some $\lambda > 0$, then $\mathbb{E}(X) = \frac{1}{\lambda}$.

(Proof)

Thus, the larger λ , the smaller the expectation.

Expected Values of Notable Distributions

Theorem

If $X \sim \text{Exp}(\lambda)$ for some $\lambda > 0$, then $\mathbb{E}(X) = \frac{1}{\lambda}$.

(Proof)

Thus, the larger λ , the smaller the expectation.

Intuition: if interarrival times between certain events are independent and $\text{Exp}(\lambda)$ -distributed, then on average λ events occur in a unit time interval. For this reason, the parameter λ of the exponential distribution is often called the **rate**.

Expected Values of Notable Distributions

Theorem

If $X \sim \text{Exp}(\lambda)$ for some $\lambda > 0$, then $\mathbb{E}(X) = \frac{1}{\lambda}$.

(Proof)

Thus, the larger λ , the smaller the expectation.

Intuition: if interarrival times between certain events are independent and $\text{Exp}(\lambda)$ -distributed, then on average λ events occur in a unit time interval. For this reason, the parameter λ of the exponential distribution is often called the **rate**.

Theorem

If X is uniformly distributed on (a, b) (where $-\infty < a < b < \infty$), then $\mathbb{E}(X) = \frac{a+b}{2}$.

(Proof)

Expectation: Positive and Negative Parts

Question: if $\int_{-\infty}^{\infty} |x|f_X(x)dx = \infty$, can we at least define $\mathbb{E}(X)$ to be $+\infty$ or $-\infty$?

The answer mirrors the discrete case:

- If $\mathbb{P}(X \geq 0) = 1$: $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf_X(x)dx \in [0, \infty]$ is well-defined.
- If $\mathbb{P}(X \leq 0) = 1$: $\mathbb{E}(X) \in [-\infty, 0]$ is well-defined.
- If X takes both positive and negative values: as before, $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$ are nonnegative random variables, and

$$X = X^+ - X^- \text{ and } |X| = X^+ + X^-.$$

If $\mathbb{E}(X^+) < \infty$ or $\mathbb{E}(X^-) < \infty$, then set

$$\mathbb{E}(X) \stackrel{\text{def.}}{=} \mathbb{E}(X^+) - \mathbb{E}(X^-),$$

which is either a real number, $+\infty$, or $-\infty$. If $\mathbb{E}(X^+) = \mathbb{E}(X^-) = \infty$, the expectation is undefined (example: Cauchy distribution).

- Again we see: $\mathbb{E}(|X|) < \infty$ iff $\mathbb{E}(X)$ exists and is finite.

Linearity of Expectation

Linearity of expectation holds not only for discrete but for *any* random variables whose expectations exist:

Theorem

Let X, Y be random variables with $\mathbb{E}(|X|) < \infty$ and $\mathbb{E}(|Y|) < \infty$. Then $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ and, for $c \in \mathbb{R}$, $\mathbb{E}(cX) = c\mathbb{E}(X)$.

The same holds, e.g., for nonnegative random variables as well, using the convention $\infty + \infty = \infty + c = \infty$ (where $c \in \mathbb{R}$):

Theorem

If X, Y are nonnegative random variables, then $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

These statements remain true even if, say, X is discrete and Y is continuous. For example, if X is the outcome of a die roll and $Y \sim \text{Exp}(1)$, then $\mathbb{E}(X + Y) = 4.5$.

