Probability and Statistics (EPRST)

Lecture 4

Review of some important probability distributions

Some distributions are so important in probability and statistics that they have their own names. We will introduce these named distributions throughout the course, starting with a very simple but useful case: an random variable that can take on only one possible value.

Definition

A random variable X has a ***one-point distribution*** *if there exists $x \in \mathbb{R}$ such that $\mathbb{P}(X = x) = 1$.*

- What is the CDF of the distribution of X?
- Give two different examples of probability spaces and random variables defined on them which have the same one-point distribution.

Two-point distributions

Definition

*A random variable X has a **two-point distribution** if there exist $x_1, x_2 \in \mathbb{R}$ ($x_1 \neq x_2$) such that*

- $\mathbb{P}(X = x_1), \mathbb{P}(X = x_2) > 0$, *and*
- $\mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) = 1$.

What is the CDF of the distribution of X? A special case:

Definition

*A random variable X has a **Bernoulli distribution** with a parameter $p \in (0, 1)$, if $\mathbb{P}(X = 1) = p = 1 - \mathbb{P}(X = 0)$.*

Binomial distribution

Definition

A random variable X has a ***binomial distribution*** *with parameters $n \in \mathbb{N}$ and $p \in (0,1)$, if*

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \; k = 0, 1, \ldots, n.$$

We will write this as $X \sim \mathrm{bin}(n, p)$. The symbol $\sim$ is read "is distributed as".

Why does this definition make sense?

Binomial distribution - cont'd

Where does this distribution appear? It is the distribution of the number of successes in n Bernoulli trials, when the probability of success in any single trial is p.
In other words, if X_i, $i = 1, \ldots, n$, denotes a random variable defined as

$$X_i = \begin{cases} 1, & \text{if a success in the } i\text{-th trial,} \\ 0, & \text{if a failure in the } i\text{-th trial,} \end{cases}$$

then the distribution of

$$X_1 + X_2 + \ldots + X_n$$

is exactly the same as the distribution of X). Observe: for every $i = 1, \ldots, n$ random variable X_i has a Bernoulli distribution with parameter p, which is a special case $X_i \sim \mathrm{bin}(1, p)$ of the binomial distribution.

Geometric distribution

Definition

A random variable X has a ***geometric distribution*** *with parameter $p \in (0,1)$ (denoted $X \sim geom(p)$), if*

$$\mathbb{P}(X = n) = (1-p)^{n-1}p, \; n = 1, 2, 3, \ldots .$$

Why does this definition make sense

Where does geometric distribution come from? It is the distribution of the waiting time (measured by the number of trials) for the first success in a sequence of Bernoulli trials with the probability of success p in any single trial.

An important property - the geometric distribution is **memoryless**: if $X \sim \mathrm{geom}(p)$, then for all $m, n \in \mathbb{N}$

$$\mathbb{P}(X > n + m | X > n) = \mathbb{P}(X > m).$$

Poisson distribution

Definition

*A random variable X has a **Poisson distribution** with parameter $\lambda > 0$ (denoted $X \sim Poiss(\lambda)$), if*

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda), \; k = 0, 1, 2, 3, \ldots .$$

Why does this definition make sense?

Poisson distribution - cont'd

Why such a distribution? It appears in a natural way as the distribution of the so called **rare events** because of the following theorem:

Theorem

If $n \to \infty$, $p_n \to 0$, *and* $np_n \to \lambda \in (0, \infty)$, *then*

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} \underset{n\to\infty}{\longrightarrow} \frac{\lambda^k}{k!} \exp(-\lambda).$$

Some important continuous distribution

We will now present a few important *continuous distributions*. Recall: these can be defined in terms of a *density function*, which takes on only non-negative values and integrates to 1.

Definition

A random variable X has a ***uniform distribution on interval*** $[a, b]$ *(denoted $X \sim U(a, b)$), if its density is*

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

Why does this definition make sense?
Why such a distibution? It serves to describe situations, in which ęach point from interval $[a, b]$ has the same chance of being picked".

Exponential distribution

Definition

*A random variable X has an **exponential distribution** with parameter $\lambda > 0$ (denoted $X \sim Exp(\lambda)$), if its density is*

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$

Why does this definition make sense?
In a sense, exponential distibution is a continuous analogue of geometric distibution. In particular exponential distribution is memoryless (like geometric): if $X \sim \mathrm{Exp}(\lambda)$, and $s, t > 0$, then

$$\mathbb{P}\left(X > t + s \middle| X > t\right) = \mathbb{P}(X > s).$$

(Standard) normal distribution

Definition

*A random variable X has a **standard normal distribution** (denoted: $X \sim \mathcal{N}(0,1)$), if its density is*

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in \mathbb{R}.$$

Why does this definition make sense?

CDF of the standard normal distribution

If $X \sim \mathcal{N}(0, 1)$, then the cumulative distribution function of X will be denoted as Φ. So $\Phi : \mathbb{R} \to \mathbb{R}$ and for all $t \in \mathbb{R}$

$$\Phi(t) = \mathbb{P}(X \leq t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x$$

Normal (or Gaussian) distribution - general case

Definition

*A random variable X has a **normal (Gaussian) distribution** with parameters $\mu \in \mathbb{R}$ and σ^2, where $\sigma > 0$ (denoted: $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$), if its density is*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$$

Why does this definion make sense?

Standardization

Observe that if

$$F_X(t) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{t} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \mathrm{d}x =$$

(substitution: $y = (x-m)/\sigma$)

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{t-m}{\sigma}} \exp\left(-\frac{y^2}{2}\right) \mathrm{d}y = \Phi\left(\frac{t-m}{\sigma}\right).$$

So: if $X \sim \mathcal{N}\left(m, \sigma^2\right)$, then

$$F_X(t) = \Phi\left(\frac{t-m}{\sigma}\right)$$

– **standardization of the distribution of** X.