STSCI 5080
Probability Models and Inference
Lecture 5: Continuous Random Variables

September 6, 2018

Binomial coefficients

For a positive integer n and $k = 0, 1, \ldots, n$,

$$\begin{aligned}\binom{n}{k} &= \text{number of } k\text{-element subsets of } \{1, \ldots, n\} \\ &= \frac{n!}{(n-k)!k!},\end{aligned}$$

where

$$n! = n(n-1)\cdots 1 \quad \text{and} \quad 0! = 1.$$

For example, $3! = 3 \cdot 2 \cdot 1 = 6, 4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, and

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{24}{2 \cdot 2} = 6.$$

PMF of $Bin(n,p)$

Theorem

The pmf of $Y \sim Bin(n,p)$ *is*

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

Proof?

Poisson random variable

Definition

Let $\lambda > 0$. X is a Possion random variable with parameter λ if its takes values in $\{0, 1, 2, \dots\}$ and its pmf is

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda},\ k = 0, 1, 2, \dots.$$

"X follows the Poisson distribution with parameter λ"

$$X \sim Po(\lambda).$$

Continuous random variable

Definition (probability density function (pdf))

A function f on $\mathbb{R}$ is a probability density function (pdf) if $f(x) \geq 0$ for any real x and

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Definition (Continuous random variable)

A random variable X is continuous if there exists a pdf f such that

$$P(X \in B) = \int_B f(x)dx$$

for any $B \subset \mathbb{R}$.

Some properties of continuous random variables

If X has pdf f, then for any fixed real x,

$$P(X = x) = \int_x^x f(y)dy = 0.$$

In addition, for any $a < b$,

$$\begin{aligned} &P(a < X < b) = P(a \leq X < b) \\ &= P(a < X \leq b) = P(a \leq X \leq b) = \int_a^b f(x)dx. \end{aligned}$$

Cumulative distribution function

Definition

Let X be a continuous random variable with pdf f. Then the cumulative distribution function (cdf) $F(x)$ of X is defined by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy$$

for any real x.

- For any $a < b$,

$$P(a < X \leq b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx$$
$$= F(b) - F(a).$$

- For $h \neq 0$,

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(y)dy.$$

So, as long as f is continuous at x, taking $h \to 0$, we have

$$F'(x) = f(x).$$

Example 5.1

Example (Uniform distribution)

Let $a < b$. A function defined by

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

is a pdf. A random variable X with this pdf is called a uniform random variable on $[a, b]$. The variable X concentrates on $[a, b]$, i.e., $P(X \in [a, b]) = 1$.

"X follows the uniform distribution on $[a, b]$"

$$X \sim U[a, b].$$

What is the cdf of X?

Example 5.2

Example

Let f be a function defined by

$$f(x) = \begin{cases} cx^2 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases},$$

where $c > 0$ is a constant. If f is a pdf, find the value of c, and compute the corresponding cdf.

Exponential random variable

Definition

Let $\lambda > 0$. A random variable X with pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is called an exponential random variable with parameter λ. The variable X concentrates on $[0, \infty)$, i.e., $P(X \in [0, \infty)) = 1$.

"X follows the exponential distribution with parameter λ"

$$X \sim Exp(\lambda).$$

What is the cdf of an exponential random variable?

Standard normal random variable

Definition

A random variable X with pdf

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \ -\infty < x < \infty,$$

is called a standard normal random variable.

"X follows the standard normal distribution"

$$X \sim N(0, 1).$$

Normal random variable with mean μ and variance σ^2

Definition

Let $-\infty < \mu < \infty, \sigma > 0$, and let $X \sim N(0,1)$. The random variable

$$Y = \mu + \sigma X$$

is called a normal random variable with mean μ and variance σ^2.

"Y follows the normal distribution with mean μ and variance σ^2"

$$Y \sim N(\mu, \sigma^2).$$

PDF of $N(\mu, \sigma^2)$

Theorem

Let $Y \sim N(\mu, \sigma^2)$. Then Y has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/(2\sigma^2)}, \ -\infty < y < \infty.$$

Proof?