Lecture 25: Oct 23

Last time

- Common Continuous Distributions

Today

- Common Continuous Distributions

Common continuous distributions

Uniform Distribution A random variable X having a pdf

$$f(x) = \begin{cases} 1 & \text{for } 0 < x \leqslant 1 \\ 0 & otherwise \end{cases}$$

is said to have a *uniform distribution* over the interval $(0, 1)$.

The cdf is:

$$F(y) = \int_{-\infty}^{y} f(x)dx = \begin{cases} 0 & \text{for } y \leqslant 0 \\ y & \text{for } 0 \leqslant y \leqslant 1 \\ 1 & \text{for } y > 1 \end{cases}$$

- Unifrom; $Y \sim U[a, b]$
- sample space: $[a, b]$
- pdf:

$$f(y) = \begin{cases} \frac{1}{b-a} & \text{for } a < y \leqslant b \\ 0 & otherwise \end{cases}$$

- cdf:

$$F(y) = \int_{-\infty}^{y} f(x)dx = \begin{cases} 0 & \text{for } y \leqslant a \\ \frac{y-a}{b-a} & \text{for } a \leqslant y \leqslant b \\ 1 & \text{for } y > b \end{cases}$$

- moments:

$$E(Y) = (a + b)/2$$
$$Var(Y) = \frac{(b-a)^2}{12}$$

Notes

- The uniform extends to the continuous case the idea of equally likely outcomes.
- If $Y \sim U[0, 1]$, then $a + (b - a)Y \sim U[a, b]$

Exponential Distribution Denoted $X \sim Exp(\lambda)$:

- sample space: $x \geqslant 0$
- pdf:

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

- cdf:

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geqslant 0 \\ 0 & \text{for } x < 0 \end{cases}$$

- moments:

$$\begin{aligned} E(X) &= 1/\lambda \\ Var(X) &= 1/\lambda^2 \\ M_X(t) &= \lambda/(\lambda - t), \quad t < \lambda \end{aligned}$$

Interpretation The exponential can be derived as the waiting time between Poisson events. Suppose that the number of events in a unit interval of time follows a Poisson(λ) distribution. Then, let Y be the time until the first event.

$$\Pr(Y > t) = \Pr(0 \text{ events in } [0, t])$$

and the number of events in $[0, t]$ follows a Poisson distribution with parameter λt. Therefore,

$$\Pr(Y > t) = e^{-\lambda t}.$$

The cdf of Y is

$$F(t) = 1 - \Pr(Y > t) = 1 - e^{-\lambda t}$$

and hence the density is $f(t) = \lambda e^{-\lambda t}$.

Alternative parameterization Many books write the density as

$$f(y) = \begin{cases} \frac{1}{\theta} e^{-y/\theta} & \text{for } y \geqslant 0 \\ 0 & otherwise \end{cases}$$

so that $E(Y) = \theta$ and $Var(Y) = \theta^2$. In this case $\theta = 1/\lambda$ is called the *mean parameter*, while $\lambda = 1/\theta$ is called the *rate parameter*.

Memoryless property The exponential has a memoryless property, just like the geometric.

$$\Pr(Y > s + t | Y > t) = \Pr(Y > s)$$

Same interpretation as the geometric for continuous time:

- The probability of an event in a time interval depends only on the length of the interval, not the absolute time of the interval.
- The underlying Poisson process is stationary: the rate λ is constant. (In the geometric case, the probability, p of getting an event in every discrete time unit is constant).

Shifted exponential Let $X \sim Exp(\lambda)$ and $Y = X + v, v \in \mathbb{R}$. Then, Y has the *shifted exponential distribution* with pdf:

$$f(y) = \begin{cases} \lambda e^{-(y-v)\lambda} & \text{for } y \geqslant v \\ 0 & otherwise \end{cases}$$

Interpretation:

- $v > 0$: Event is delayed
- $v < 0$: The news of the event is delayed

Does the shifted expoenential maintain the memoryless property?

Double exponential The *double exponential distribution* is formed by reflecting an exponential distribution around zero. It has pdf:

$$f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}, \quad x \in \mathbb{R}$$

Laplace exponential Suppose X has the above distribution with $\lambda = 1$. Now let $Y = \sigma X + \mu, \mu \in \mathbb{R}$ (shifting) and $\sigma > 0$ (scaling). Then Y has the *Laplace distribution* with pdf:

$$f_Y(y) = \frac{1}{2\sigma} \exp\left(-\frac{|y-\mu|}{\sigma}\right)$$

with moments

$$EY = \mu, \quad Var(Y) = 2\sigma^2$$

The Laplace distribution provides an alternative to the normal for centered data with fatter tails but all finite moments.

Normal Distribution Introduced by De Moivre (1667 - 1754) in 1733 as an approximation to the binomial. Later studied by Laplace and others as part of the Central Limit Theorem. Gauss derived the normal as a suitable distribution for outcomes that could be thought of as sums of many small deviations.

- Sample space: $\mathbb{R} = (-\infty, \infty)$
- pdf: For $Y \sim N(\mu, \sigma^2)$,

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \quad -\infty < y < \infty$$

- cdf: There is no closed form.
- When $\mu = 0$ and $\sigma = 1$, the distribution is called *standard normal*:

$$\Phi(y) = \Pr(Y \leqslant y), \quad \Phi(-y) = 1 - \Phi(y)$$

- Mean:

$$EY = \mu$$

- Variance:

$$Var(Y) = E(Y-\mu)^2 = \sigma^2$$

- Higher central moments:

$$E(Y-\mu)^m = \begin{cases} \frac{m!}{2^{m/2}(m/2)!}\sigma^m & m \text{ is even} \\ 0 & m \text{ is odd} \end{cases}$$

- In particular:

$$\mu_3 = E(Y-\mu)^3 = 0(\text{Skewness})$$
$$\mu_4 = E(Y-\mu)^4 = 3\sigma^4$$

- Moment generating function:

$$M_Y(t) = \exp(\mu t + \sigma^2 t^2/2)$$

Standardization

$$Y \sim N(\mu, \sigma^2) \iff Z = \frac{Y-\mu}{\sigma} \sim N(0,1)$$

Shifting and scaling:

$$Z \sim N(0,1) \iff Y = \sigma Z + \mu \sim N(\mu, \sigma^2)$$

Notes

- Normal distribution is useful in many practical settings. E.g. measurement error.
- Plays an important role in *sampling distributions* in *large samples*, since the Central Limit Theorem syas that the sums of independent identically distributed random variables are approximately normal
- There are many important distributions that can be derived from functions of normal random variables (e.g. χ^2, t, F). We will briefly present the pdf's and sample spaces of these distributions.