Review (2.6)

- Poisson Processes, Poisson distribution.
- Mean, variance and m.g.f. of Poisson distribution.
- Poisson Approximation to Binomial distribution.

Today's Lecture (3.1,3.3)

- Continuous random variable and its c.d.f.
- Probability density function (p.d.f.), its properties.
- Mean, variance, moment generating function
- Percentiles for p.d.f.

Continuous-type random variable and its c.d.f.

The random variable X is a continuous random variable if X takes all values in an interval of numbers.

- Wheel of fortune: the angle of the pointer is within $[0^o, 360^o)$.
- The length of time it takes to check out at Walmart in the weekend.
- The time we wait to see the next volcano eruption in the world.

Note:

- When X is continuous, $P(X = x) = 0$ for all x. The probability mass function is meaningless.

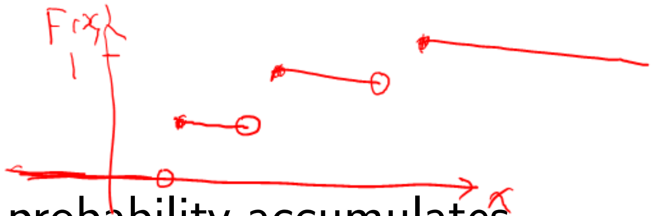
- Although we cannot assign a probability to any value of X, we are able to assign probabilities to intervals: e.g. $P(X = 1) = 0$, but $P(0.999 \leq X \leq 1.001)$ can be > 0.

Comparison between discrete and continuous r.v.

Same:

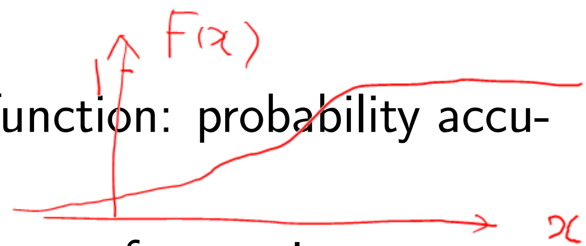
- $F_X(x) = P(X \leq x)$.
- $P(a < X \leq b) = F(b) - F(a)$.

Different:



- Discrete r.v.
  - $F_X(x)$ is a step function: probability accumulates in discrete steps.
  - Endpoints are very important for discrete r.v. For a discrete r.v. taking integer values, if a, b are integers, then

$$\begin{aligned} P(a < X < b) &= P(a+1 \leq X \leq b-1) \\ &= F_X(b-1) - F_X(a) \end{aligned}$$



- Continuous r.v.
  - $F_X(x)$ is a continuous function: probability accumulates continuously.
  - Endpoints are not important for continuous r.v.:

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X < b) = P(a \leq X < b) \\ &= P(a < X \leq b) = F(b) - F(a) \end{aligned}$$

Probability density function

Definition: The probability density function (p.d.f.) of a continuous random variable X is

$$f_X(x) = \lim_{t \to 0} \frac{F_X(x+t) - F_X(x)}{t} = F'_X(x)$$

Properties: The probability density function (p.d.f.) of a random variable X of the continuous type, with space S that is an interval or union of intervals, is an integrable function $f(x)$ satisfying the following conditions:

- $f(x) > 0$, $x \in S$.
- $\int_S f(x)dx = 1$
- If $(a, b) \subset S$, the probability of the event $\{a < X < b\}$ is
$$P(a < X < b) = \int_a^b f(x)dx$$

Example: $f(x) = 2x$, $0 \le x \le 1$.

- What is S? $S = \{x : 0 \le x \le 1\} = [0, 1]$
- What is $P(1/4 \le X \le 3/4)$?

$$= \int_{\frac{1}{4}}^{\frac{3}{4}} 2x \, dx = x^2 \Big|_{\frac{1}{4}}^{\frac{3}{4}} = \frac{8}{16} = \frac{1}{2}$$

Example

- We can extend the definition of p.d.f. $f_X(x)$ to the entire set of real numbers by letting it equal zero when $x \notin S$.
- $F_X(x) = \int_{-\infty}^{x} f_X(u)du$

Example.

Let $f_X(x) = \begin{cases} ce^{-2x} & \text{for } 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}$

exponential distr'n

- Find the constant c.

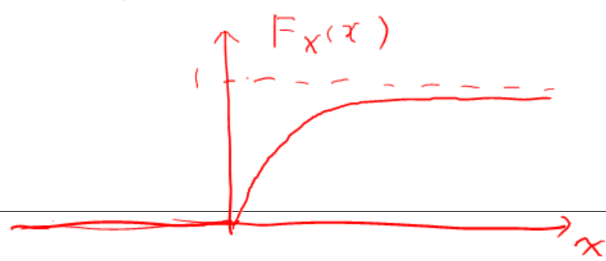
$$1 = \int_{-\infty}^{+\infty} f_X(x)dx = \int_0^{+\infty} c e^{-2x} dx = -\frac{c}{2} e^{-2x} \Big|_0^{+\infty} = \frac{c}{2}$$

$$\Rightarrow c = 2$$

- Find the cumulative distribution function $F_X(x)$ for all x.

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du = \begin{cases} 0 & \text{if } x < 0 \\ \int_0^x 2e^{-2u}du = 1 - e^{-2x} & x \geq 0 \end{cases}$$

as we can see, as $x \to +\infty$, $F_X(x) \to 1$



Mean, variance and Moment generating function

The expected value (mean) of X is

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

The variance of X is

$$\sigma^2 = \mathrm{Var}(X) = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$

OR, $= EX^2 - \mu^2 = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - \mu^2$

The standard deviation of X is

$$\sigma = \sqrt{\mathrm{Var}(X)}$$

The moment-generating function, if it exists, is

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx, \quad -h < t < h$$

If $M(t)$ exists on an open interval that contains $\{0\}$, then $M(t)$ determines the distn.

The following results still hold for continuous r.v.

$$\begin{aligned} \sigma^2 &= E(X^2) - \mu^2 \\ \mu &= M'(0) \\ \sigma^2 &= M''(0) - [M'(0)]^2 \end{aligned}$$

- The definitions associated with mathematical expectation are the same as those in the discrete case except that integrals replace summations.

Example

Let X has the p.d.f.

$$f(x) = \begin{cases} 5e^{-5x}, & 0 \le x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

- $E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx = \int_0^{+\infty} 5xe^{-5x}\,dx = \cdots$

  use integration by part

- $M(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) = 5\int_0^{+\infty} e^{(t-5)x}\,dx$

  $= \frac{5}{t-5} e^{(t-5)x} \Big|_0^{+\infty}$

  $= \frac{5}{t-5}(0-1)$ if $(t-5)<0$

  $= \frac{1}{1-t/5}$; $t<5$, or $t \in (-\infty, 5)$

$$\begin{aligned}
M'(t) &= \frac{1}{5(1-t/5)^2}, \quad t<5 \\
M''(t) &= \frac{2}{25(1-t/5)^3}, \quad t<5 \\
\mu &= M'(0) = 1/5 \\
\sigma^2 &= M''(0) - [M'(0)]^2 = 2/25 - (1/5)^2 = 1/25
\end{aligned}$$

Percentiles for p.d.f.

- Given a p.d.f. $f(x)$ for a random variable X.
- $100p$th percentile is a number π_p such that

$$p = \int_{-\infty}^{\pi_p} f(x)dx = F(\pi_p)$$

Illustration:

- median $m = \pi_{0.50}$: 50th percentile.
- first (third) quartile: $q_1 = \pi_{0.25}$, $q_3 = \pi_{0.75}$.
- quantile of order p: the $100p$th percentile.

Is this a valid density?

Example: Let X has p.d.f. $f(x) = x/4$, $1 < x < 3$,

- median : m

$$\frac{1}{2} = \int_{-\infty}^{m} f(x)\,dx = \int_{1}^{m} \frac{x}{4}\,dx = \frac{x^2}{8}\Big|_1^m = \frac{m^2-1}{8}$$

$$\Rightarrow m = \sqrt{5}$$

- $\pi_{0.90}$

$$0.9 = \int_{-\infty}^{\pi_{0.9}} f(x)\,dx = \int_{1}^{\pi_{0.9}} \frac{x}{4}\,dx = \frac{\pi_{0.9}^2 - 1}{8}$$

$$\Rightarrow \pi_{0.9} = \sqrt{8.2}$$