Continuous Random Variables and Probability Distributions

Definition$X$ is a continuous random variable if its possible values for an interval. For example, $A < x < B$.
pmf is not proper for continuous r.v.
A physical analog: probability = mass
Limit of density histogram

Definition:(pdf)
Let X be a continuous r.v.. Then a probability distribution or **probability density function (pdf)** is a function $f(x)$ such that

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

The graph of $f(x)$ is called the density curve

Two conditions
1.
2.

$P(a \leq X \leq b)$; $P(a < X \leq b)$; $P(a \leq X < b)$; $P(a < X < b)$

Two simple example
1. Uniform Distribution
X: waiting time for bus

$$f(x) = \begin{cases} 1/5 & 0 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$$

2. Exponential distribution:
X: the time between two consecutive phone calls

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

Examples 4.3 and 4.4

Cumulative distribution function

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

Example Uniform distribution over $[A, B]$

Calculate probability using cdf

$$P(X > a) = 1 - F(a)$$
$$P(a \le X \le b) = F(b) - F(a)$$

Obtaining f(x) from F(x)

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

(100p)th percentile $\eta(p)$

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

Median

Expected value or mean

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

$$E(h(X)) = \mu_{h(X)} = \int_{-\infty}^{+\infty} h(x) \cdot f(x)dx$$

$$E(aX + b) = aE(X) + b$$

Variance and standard deviation

$$\sigma_X^2 = V(X) = E((X - \mu)^2) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx$$

$$\sigma_X = \sqrt{V(X)}$$

Shortcut for variance calculation

$$V(X) = E(X^2) - (E(X))^2$$

$$V(aX+b) = a^2 V(X)$$

Please read all the examples in section 4.2 carefully

Normal Distributions

Definition A continuous r.v. X is said to have a normal distribution with parameters μ and σ if the pdf of X is

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

where $-\infty < \mu + \infty$ and $\sigma > 0$.
0. $X \sim N(\mu, \sigma)$
1.

$$\int_{-\infty}^{+\infty} f(x;\mu,\sigma)dx = 1$$

2. $E(X) = \mu$
3. $V(X) = \sigma^2$
4. Density curve

Standard Normal Distribution

Definition The normal distribution with parameters $\mu = 0$ and $\sigma = 1$ is called a standard normal distribution. A random variable with a standard normal distribution is called a standard normal random variable and will be denoted by Z.

The pdf of Z:

$$f(z;0,1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

The cdf of Z:

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

Cumulative probabilities for standard normal distribution

Appendix Table A.3

Calculate probability for Z using Table A.3

1. $P(Z \leq 1.25)$
2. $P(Z > 1.25)$
3. $P(Z \leq -1.25)$
4. $P(-.38 \leq Z \leq 1.25)$

How about nonstandard normal distribution? **Proposition**

If $X \sim N(\mu, \sigma)$, then $\frac{X-\mu}{\sigma}$ has a standard normal distribution, i.e.,

$$Z = \frac{X - \mu}{\sigma}$$

Thus,

$$P(a \leq X \leq b) = P(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma})$$

$$P(X \leq a) = P(Z \leq \frac{a-\mu}{\sigma})$$

$$P(X \geq b) = P(Z \geq \frac{b-\mu}{\sigma}$$

Example: $X \sim N(1.25, 0.46)$

1. $P(1.00 \leq X \leq 1.75)$
2. $P(X > 2)$

Percentiles of normal distribution

Standard normal distribution

$100p$th percentile

10th percentile:

60th percentile:

97th percentile:

z_α notation

Upper tail with area α.

$100(1-\alpha)$ percentile of the standard normal distribution

(critical values)

Important standard normal percentils and critical values

p	α	z_α
0.9	0.1	1.28
0.95	0.05	1.645
0.975	0.025	1.96
0.99	0.01	2.33
0.995	0.005	2.58

General normal distribution $N(\mu, \alpha)$
100pth percentile of $X \sim N(\mu, \alpha)$:

$$p = P(X \leq x) = P(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}) = P(Z \leq \frac{x-\mu}{\sigma})$$

$\frac{x-\mu}{\sigma}$ is the 100pth percentile of the standard normal distribution that can be derived from Table A.3

$$\begin{array}{c}(100p)\text{th percentile}\\ \text{for } N(\mu,\sigma)\end{array} = \mu + \left(\begin{array}{c}(100\text{p})\text{th percentile}\\ \text{for standard normal}\end{array}\right) \cdot \sigma$$

Example: find the 60th percentile of $N(2, 4)$

Normal approximation to binomial distribution Proposition Let X be a binomial rv based on n trials with success probability p. Then if the binomial probability histogram is not too skewed, X has approximately a normal distribution with $\mu = np$

and $\sigma = \sqrt{npq}$. In particular, for $x =$ a possible value of X,

$$P(X \leq x) = B(x; n, p) \approx \begin{pmatrix} \text{area under the normal curve} \\ \text{to the left of } x + .5 \end{pmatrix}$$

$$= \Phi(\frac{x + 0.5 - np}{\sqrt{npq}})$$

In practicem the approximation is adequate provided that both $np \geq 10$ and $nq \geq 10$

Example 4.19

Exponential Distributions

$$X \sim \text{Exp}(\lambda)$$

pdf:

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

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

cdf

$$F(x; \lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Applications

elapsed time between the occurence of two successive events
life time of electronic device

Chi-square Distributions

$$X \sim \chi^2(\nu): \ \nu \text{ is the number of degrees of freedom}$$

pdf

$$f(x;\nu) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

0. Gamma function:

$$\Gamma(\nu/2) = \int_0^{+\infty} x^{\nu/2-1} e^{-x} dx$$

1.

$$E(X) = \nu$$

2.

$$V(X) = 2\nu$$

cdf:

$$F(x;\nu) = \begin{cases} \int_0^x f(y;\nu)dy & x \geq 0 \\ 0 & x < 0 \end{cases}$$

critical values are give in Table A.7

Density Curve

Application

Used in statistical inference