

Probability Theory

Chapter 5—Continuous random variables

Nan Liu¹

¹These slides are based on Prof. Mui Li's. Any mistakes are my own. If you find an error, please send an email at nanliu@xmu.edu.cn

5.1 Introduction: Continuous Random Variables

- For the time being, we have only considered discrete random variables (r.v.) - set of possible values is finite or countable - such as the number of arrivals in a given time instants, the number of successful trials, the number of items with a given characteristic.
- In many practical applications, we have to deal with random variables whose set of possible values is uncountable.
- **Example: Waiting for a bus.** The time (in minutes) which elapses between arriving at a bus stop and a bus arriving can be modelled as a r.v. X taking values in $[0, \infty)$.
- **Example: Share price.** The values of one share of a specific stock at some given future time can be modelled as a r.v. X taking values in $[0, \infty)$.
- **Example: Weight.** The weight of a randomly chosen individual can be modelled as a r.v. X taking values in $[0, \infty)$.
- **Example: Temperature.** The temperature in Xiamen can be modelled as a r.v. X taking values in $[-10, +40]$ degrees celsius.

5.1 Introduction: Continuous Random Variable

- **'Formal' definition:** We say that X is a (real-valued) continuous r.v. if there exists a nonnegative function $f : \mathbb{R} \rightarrow [0, \infty)$ such that for any set A of real numbers

$$P(X \in A) = \int_A f(x) dx.$$

- $f(x)$ is called the *probability density function* (pdf) of the r.v. X and the associated (*cumulative*) *distribution function* is

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

so we have

$$f(x) = \frac{dF(x)}{dx}.$$

- We also have

$$P(X \in \mathbb{R}) = 1 = \int_{\mathbb{R}} f(x) dx$$

5.1 Introduction: Continuous Random Variable

- Assume we have a set $A = [a, b]$ then

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

- $P(X \in [a, b])$ corresponds to the area under the curve $f(x)$ between a and b .
- Consider a small interval $[x, x + \Delta x]$, then we have

$$P(x \leq X \leq x + \Delta x) = \int_x^{x+\Delta x} f(y) dy \approx f(x) \Delta x$$

so $P(x \leq X \leq x + \Delta x)$ is *proportional* to $f(x)$ for Δx small.

5.1 Important Remarks

- $f(x)$ is NOT a probability, e.g. $f(1)$ is not the probability that $X = 1$, it is a probability density.
- As $f(x)$ is NOT a probability, it is NOT surprising to have $f(x) > 1$ for some values of $x \in \mathbb{R}$.
- For a continuous r.v., the probability of any single value is zero; i.e. for any $x \in \mathbb{R}$ we have $P(X = x) = 0$. For example if you model a waiting time X by an exponential density, then the proba of waiting 5,123123123123... or exactly 5.0 minutes will be zero!
- $P(X = x) = 0$ for any x implies that $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$.
- Keep track of the “range” of a density, such as the interval $[0, 1]$ or $[0, \infty)$ when doing your calculations.

5.1 Toy example

- Assume you are told that the pdf of a r.v. X satisfies

$$f(x) = \begin{cases} Cx & \text{if } x \in [0, 1] \\ Cx^2 & \text{if } x \in [-1, 0] \\ 0 & \text{otherwise.} \end{cases}$$

where C is an unknown constant. What is the probability that $X < 0.5$.

- We first determine the value of C using the fact that $\int_{\mathbb{R}} f(x) dx = 1$;

$$\begin{aligned} \int_{\mathbb{R}} f(x) dx &= C \int_{-1}^0 x^2 dx + C \int_0^1 x dx \\ &= \frac{C}{3} + \frac{C}{2} = \frac{5C}{6} = 1 \Rightarrow C = \frac{6}{5}. \end{aligned}$$

- The probability that $X < 0.5$ is

$$\Pr(X < 0.5) = \frac{6}{5} \left[\int_{-1}^0 x^2 dx + \int_0^{0.5} x dx \right] = \frac{6}{5} \left(\frac{1}{3} + \frac{1}{8} \right)$$

5.1 Example: Insurance Policy

- A group insurance policy covers the medical claims of the employees of a small company. The value, V , of the claims made in one year is described by

$$V = 100,000X$$

where X is a random variable with pdf

$$f(x) = \begin{cases} c(1-x)^4 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

- What is the conditional probability that V exceeds 40,000 given that V exceeds 10,000?

5.1 Example: Insurance Policy

- We are interested in

$$\begin{aligned}P(V > 40,000 | V > 10,000) &= \frac{P(V > 40,000 \cap V > 10,000)}{P(V > 10,000)} \\&= \frac{P(V > 40,000)}{P(V > 10,000)}\end{aligned}$$

where

$$P(V > v) = P(100,000X > v) = P\left(X > \frac{v}{100,000}\right)$$

- First we need to determine c using $\int_0^1 f(x) dx = 1$; that is

$$\begin{aligned}\int_0^1 f(x) dx &= c \left[-\frac{(1-u)^5}{5} \right]_0^1 = \frac{c}{5} \\&\Rightarrow c = 5.\end{aligned}$$

5.1 Example: Insurance Policy

- We need to compute the cdf $F_X(x)$ of X which is given by

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ c \left[-\frac{(1-u)^5}{5} \right]_0^x = 1 - (1-x)^5 & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$$

- So we are interested in

$$P(V > 40,000 | V > 10,000) = \frac{1 - F_X(0.4)}{1 - F_X(0.1)} = \frac{0.078}{0.590} = 0.132.$$

5.1 Introduction: Examples of special distributions

- **Uniform density.** Consider a r.v. X with pdf given for $c < d$

$$f(x) = \begin{cases} \frac{1}{d-c} & \text{if } x \in [c, d] \\ 0 & \text{otherwise} \end{cases}$$

- **Exponential density.** Consider a r.v. X with pdf given for $\lambda > 0$ by

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x \in [0, \infty) \\ 0 & \text{otherwise} \end{cases}$$

- **Normal density.** Consider a r.v. X with pdf given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty, \quad 0 < \sigma < \infty$$

5.2 (i) Expectation of Continuous Random Variables

- We define the expected value of an r.v. X by

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

- More generally for any real-valued function $g : \mathbb{R} \rightarrow \mathbb{R}$ then

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

5.2 (i) Expectation of Continuous Random Variables

- **Uniform density.** We have for $c < d$ and $x \in [c, d]$

$$f(x) = \frac{1}{d - c}$$

then

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{d - c} \int_c^d x dx = \frac{d^2 - c^2}{2(d - c)} = \frac{c + d}{2}.$$

- **Exponential density.** We have for $\lambda > 0$ and $x \geq 0$

$$f(x) = \lambda \exp(-\lambda x)$$

so

$$\begin{aligned} E(X) &= \lambda \int_0^{\infty} x \exp(-\lambda x) dx \\ &= \lambda \left[x \frac{\exp(-\lambda x)}{-\lambda} \right]_0^{\infty} - \lambda \int_0^{\infty} \frac{\exp(-\lambda x)}{-\lambda} dx \\ &= \frac{1}{\lambda}. \end{aligned}$$

5.2 (i) Expectation of Continuous Random Variables

- **Even density.** For $f(x) = f(-x)$, we have

$$\begin{aligned} E(X) &= \int_{-\infty}^0 x f(x) dx + \int_0^{\infty} x f(x) dx \\ &= \int_{-\infty}^0 x f(-x) dx + \int_0^{\infty} x f(x) dx \\ &= - \int_0^{\infty} u f(u) dx + \int_0^{\infty} x f(x) dx \\ &= 0 \end{aligned}$$

5.2 (i) Expectation of Continuous Random Variables

- **Cauchy density.** Consider the pdf

$$f(x) = \frac{1}{(x+1)^2} \text{ for } x \geq 0$$

then

$$\begin{aligned} E(X+1) &= \int_0^{\infty} \frac{x+1}{(x+1)^2} dx = \int_0^{\infty} \frac{1}{x+1} dx \\ &= \lim_{u \rightarrow \infty} [\log(x+1)]_0^u = \infty \end{aligned}$$

Hence we can conclude that $E(X)$ is infinite in this case ($E[X]$ does not exist).

- Distributions such that $E(X)$ is not finite are sometimes referred to as heavy-tails; they appear a lot in finance, actuarial science etc.

5.2 (i) Example: Selling Printers

- The lifetime of a printer costing 200\$ is exponentially distributed with mean 2 years. The manufacturer agrees to pay a full refund to a buyer if the printer fails during the first year following its purchase, and a one-half refund if it fails during the second year. If the manufacturer sells 100 printers, how much should it expect to pay in refunds?
- Let T denote a printer lifetime then

$$f(t) = \frac{1}{2} \exp\left(-\frac{t}{2}\right) \mathbf{1}_{(0,\infty)}(t)$$

so we have

$$\begin{aligned} P(T < 1) &= \int_0^1 f(t) dt = [\exp(-t/2)]_0^1 \\ &= 1 - \exp(-1/2) = 0.393, \\ P(1 < T < 2) &= \int_1^2 f(t) dt = [\exp(-t/2)]_1^2 \\ &= \exp(-1/2) - \exp(-1) = 0.239. \end{aligned}$$

5.2 (i) Example: Selling Printers

- Let X_i denote the refund associated to the i th printer sold. Then for any $i = 1, \dots, 100$

$$X_i = \begin{cases} 200 & \text{with proba } 0.393 \\ 100 & \text{with proba } 0.239 \\ 0 & \text{with proba } 0.368 \end{cases}$$

so we have

$$E(X_i) = 200 \times 0.393 + 100 \times 0.239 = 102.56.$$

- The expected refund associated to the 100 printers sold is thus

$$E\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} E(X_i) = 100 \times 102.56 = 10,256.$$

5.2 (i) Example: Failure Discovery

- A device that continuously measures and records seismic activity is placed in a remote region. The time, T , to failure of this device is exponentially distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T, 2)$. What is the expected value of X ?
- We use the formula $E(g(T)) = \int g(t) f(t) dt$ for $f(t)$ an exponential of parameter $1/3$ and $g(t) = \max(t, 2)$ so

$$\begin{aligned} E(X) &= \int_0^{\infty} \max(t, 2) \frac{1}{3} \exp\left(-\frac{t}{3}\right) dt \\ &= \int_0^2 \frac{2}{3} \exp\left(-\frac{t}{3}\right) dt + \int_2^{\infty} \frac{t}{3} \exp\left(-\frac{t}{3}\right) dt \\ &= \left[-2 \exp\left(-\frac{t}{3}\right)\right]_0^2 - \left[t \exp\left(-\frac{t}{3}\right)\right]_2^{\infty} + \int_2^{\infty} \frac{1}{3} \exp\left(-\frac{t}{3}\right) dt \\ &= -2 \exp\left(-\frac{2}{3}\right) + 2 + 2 \exp\left(-\frac{2}{3}\right) - \left[3 \exp\left(-\frac{t}{3}\right)\right]_2^{\infty} \\ &= 2 + 3 \exp\left(-\frac{2}{3}\right) = 3.54 \end{aligned}$$

5.2 (i) Expectation

- **Question:** Assume $E|X| < \infty$. Show that

$$E[X] = \arg_a \operatorname{Min} E[(X - a)^2]$$

5.2 (ii) Variance of Continuous Random Variables

- We define the variance as

$$\begin{aligned}\text{Var} [X] &= E \left((X - E(X))^2 \right) \\ &= E(X^2) - E(X)^2\end{aligned}$$

5.2 (ii) Variance of Continuous Random Variables

- **Uniform density.** We have for $f(x) = \frac{1}{d-c}$ for $x \in [c, d]$ and $E(X) = \frac{c+d}{2}$. We also have

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \frac{1}{d-c} \int_c^d x^2 dx = \frac{d^3 - c^3}{3(d-c)} \\ &= \frac{c^2 + d^2 + cd}{3} \end{aligned}$$

so

$$\begin{aligned} \text{Var}[X] &= \frac{c^2 + d^2 + cd}{3} - \frac{(c+d)^2}{4} \\ &= \frac{(d-c)^2}{12} \end{aligned}$$

5.2 (ii) Example: Repair Cost and Insurance Payment

- The owner of an automobile insures it against damage by purchasing an insurance policy with a deductible of 250\$. In the event that the automobile is damaged, repair costs can be modeled by a uniform random variables on the interval $(0, 1500)$. Determine the standard deviation of the insurance payment in the event that the automobile is damaged.
- Let X be the repair cost and Y the insurance payment then

$$Y = \begin{cases} 0 & \text{if } X < 250 \\ X - 250 & \text{if } X \geq 250 \end{cases}$$

and we want to compute $\sqrt{\text{Var}(Y)}$.

5.2 (ii) Example: Repair Cost and Insurance Payment

- We have

$$E(Y) = \int_{250}^{1500} \frac{1}{1500} (x - 250) dx = \frac{1}{3000} \left[(x - 250)^2 \right]_{250}^{1500} = 521,$$

$$E(Y^2) = \int_{250}^{1500} \frac{1}{1500} (x - 250)^2 dx = \frac{1}{4500} \left[(x - 250)^3 \right]_{250}^{1500} = 434,028.$$

- Finally, we obtain

$$\begin{aligned} \text{Var}(X) &= E(Y^2) - E(Y)^2 = 434,028 - 521^2, \\ \sqrt{\text{Var}(Y)} &= 403. \end{aligned}$$

5.2 (ii) Variance of Continuous Random Variables

- **Exponential density.** We have $f(x) = \lambda \exp(-\lambda x)$ for $\lambda > 0$ and $x \geq 0$ and $E(X) = \frac{1}{\lambda}$. We have

$$\begin{aligned} E(X^2) &= \lambda \int_0^{\infty} x^2 \exp(-\lambda x) dx \\ &= \lambda \left[x^2 \frac{\exp(-\lambda x)}{-\lambda} \right]_0^{\infty} - \lambda \int_0^{\infty} 2x \frac{\exp(-\lambda x)}{-\lambda} dx \\ &= \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2} \end{aligned}$$

so

$$\text{Var}(X) = E(X^2) - E(X)^2 = \frac{1}{\lambda^2}.$$

5.2 (iii) Quantile of Continuous Random Variables: Median

- The median of a continuous r.v. X of pdf $f(x)$ is the number m such that

$$\int_{-\infty}^m f(x) dx = \int_m^{\infty} f(x) dx = \frac{1}{2};$$

that is the number m such that

$$\Pr(X \leq m) = P(X \geq m) = \frac{1}{2}.$$

- For example, assume we look at a population of people. Let X be the salary of a randomly chosen person from this population of pdf $f(x)$, and let m be the median salary of the population. This means that half the population earns less than m dollars and half earns more than m dollars.

5.2 (iii) Quantile of Continuous Random Variables: Median

- **Uniform density.** For $c < d$, we have $f(x) = \frac{1}{d-c}$ and the median is $m = \frac{c+d}{2}$; i.e. in this case the median and $E(X)$ are similar.
- **Exponential density.** We have $f(x) = \lambda \exp(-\lambda x)$ for $\lambda > 0$ and $x \geq 0$.
- The median corresponds to the value

$$\int_0^m \lambda \exp(-\lambda x) dx = \int_m^\infty \lambda \exp(-\lambda x) dx = \frac{1}{2}.$$

- We have

$$\begin{aligned} \int_m^\infty \lambda \exp(-\lambda x) dx &= [\exp(-\lambda x)]_m^\infty \\ &= \exp(-\lambda m) \end{aligned}$$

and

$$\exp(-\lambda m) = \frac{1}{2} \Leftrightarrow m = \frac{\log 2}{\lambda}$$

- In this case, the median and $E(X)$ are different.

5.2 (iii) Quantile of Continuous Random Variables: Median

- **Difference Between Mean and Median:**

- Median is the cutoff point that divides the population in half.
- Mean can be misleading when used to measure the location of highly skewed data. In contrast, median is a more robust measure of the central tendency of a distribution in the sense that it is not much affected by a few outliers.
- Median $Q(0.5)$ is the optimal solution for minimizing the *mean absolute error*, that is,

$$Q(0.5) = \arg \min_a \{E|X - a|\},$$

while mean is the optimal solution for minimizing the *mean square error*, that is,

$$E(X) = \arg \min_a \{E(X - a)^2\}.$$

5.2 (iii) Quantile of Continuous Random Variables: Median

- **Example:** For symmetric distribution, e.g., normal distribution, mean and median are the same. For skewed distributions, mean and median are different.

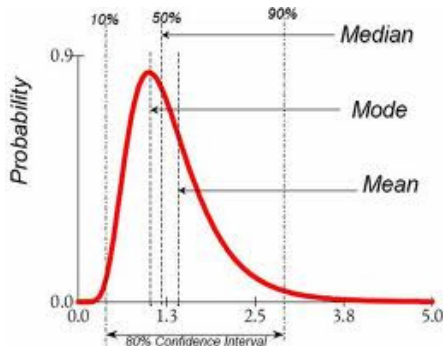


Figure: Mean and Median.

5.2 (iii) Quantile of Continuous Random Variables

- **Definition: α -Quantile.** Suppose r.v. X has a continuous cdf $F_X(x)$. Then the α -quantile ($0 \leq \alpha \leq 1$) of the distribution $F_X(x)$ is defined as $Q(\alpha)$, which satisfies

$$F_X(Q(\alpha)) = P(X \leq Q(\alpha)) = \alpha.$$

- **Remarks:**

- For α -quantile $Q(\alpha)$, we have

$$\int_{-\infty}^{Q(\alpha)} f_X(x) dx = \alpha.$$

- 0.5-quantile is called the *median*, 0.25-quantile and 0.75-quantile are called *lower* and *upper quartiles*.
- 0-quantile is the minimum value, 1-quantile is the maximum value.

5.2 (iii) Quantile of Continuous Random Variables

- **Boxplot:**

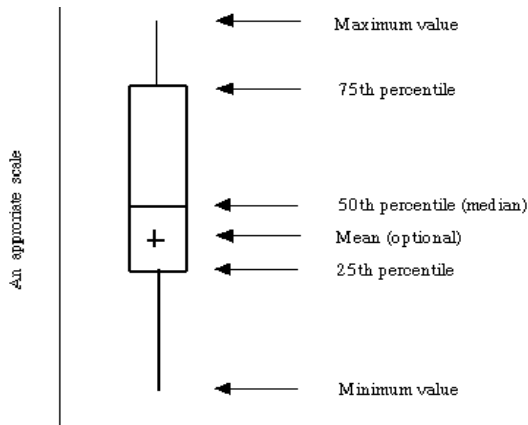


Figure: Boxplot.

5.3: Uniform Variable

- Refer to previous slides.

5.4: Normal Random Variables

- Also known as Gaussian random variables in the literature.
- We say that X is a normal r.v. of parameters (μ, σ^2) if its pdf is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty.$$

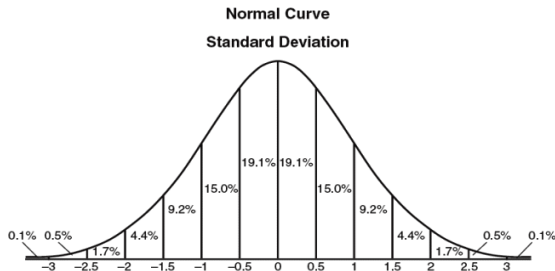


Figure: PDF of standard normal distribution.

5.4 Normal Random Variables

- The normal distribution is often used to describe, at least approximately, any variable that tends to cluster around the mean; e.g. the heights of USA males are roughly normally distributed. A histogram of male heights will appear similar to a bell curve, with the correspondence becoming closer if more data are used.

5.4 Properties of Normal Random Variables

- It can indeed be checked that

$$\int_{-\infty}^{\infty} \exp \left(-\frac{(x - \mu)^2}{2\sigma^2} \right) dx = \sqrt{2\pi}\sigma.$$

- We have also

$$E(X) = \mu$$

and

$$\text{Var}(X) = \sigma^2$$

- Hence μ is referred to as the mean and σ^2 as the variance.

5.4 The 68-95-99.7 Rule

- We have

$$P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68,$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95,$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.997.$$

- This helps doing quickly some approximate calculations.
- The distribution of the scores of the more than 1.3 million high school seniors in 2002 who took the SAT verbal exam is close to normal with $(\mu, \sigma^2) = (504, 111^2)$.
- Hence 95% of the SAT scores are between $504 - 222 = 282$ and $504 + 222 = 726$. The other 5% of scores lie outside this range.

5.4 Cumulative Distribution Function

- Consider Z a normal r.v. of parameters $(\mu = 0, \sigma^2 = 1)$; known as *standard* r.v. in the literature.
- It is customary to denote $\Phi(z)$ the cdf of Z ; i.e.

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{(y-\mu)^2}{2}\right) dy.$$

- $\Phi(z)$ does not admit an analytical expression but is tabulated for $z \geq 0$.
- One can easily show that

$$\Phi(-z) = P(Z \leq -z) = P(Z \geq z) = 1 - \Phi(z)$$

5.4 Standardizing normal variables

- Let X a normal r.v. of mean μ and variance σ^2 .
- Define the new r.v.

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

then Z is a standard normal r.v.

- Hence

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right). \end{aligned}$$

5.4 Properties of Normal Random Variables

- Let X be a normal r.v. of parameters (μ, σ^2) and consider the new r.v. Y such that

$$Y = aX + b$$

then we know that

$$\begin{aligned} E(Y) &= aE(X) + b = a\mu + b, \\ \text{Var}(Y) &= a^2 \text{Var}(X) = a^2 \sigma^2. \end{aligned}$$

- A much stronger result is true, Y is a normal r.v. of parameters $(a\mu + b, a^2\sigma^2)$ as

$$P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$

- The *chain rule* tells us that $[u(v(y))]' = v'(y) \cdot u'(v(y))$ so

$$\begin{aligned} f_Y(y) &= \frac{1}{a} f_X\left(\frac{y-b}{a}\right) \\ &= \frac{1}{a} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma a} \exp\left(-\frac{(y-b-a\mu)^2}{2\sigma^2 a^2}\right) \end{aligned}$$

5.4 Example

- Let X a normal r.v. of mean $\mu = 2$ and variance $\sigma^2 = 25$. Assume you want to compute using the table of $\Phi(x)$ (a) $P(1 \leq X \leq 4)$, (b) $P(X > 0)$ and (c) $P((X - 2)^2 > 5)$
- (a) We have

$$\begin{aligned}P(1 \leq X \leq 4) &= P\left(\frac{1-2}{5} \leq \frac{X-2}{5} \leq \frac{4-2}{5}\right) \\&= P\left(\frac{-1}{5} \leq Z \leq \frac{2}{5}\right) = \Phi\left(\frac{2}{5}\right) - \Phi\left(-\frac{1}{5}\right) \\&= \Phi\left(\frac{2}{5}\right) - \left(1 - \Phi\left(\frac{1}{5}\right)\right)\end{aligned}$$

where Z is a normal r.v. of mean 0 and variance 1; i.e. a standard normal r.v.

5.4 Example

- (b) We have

$$\begin{aligned}P(X > 0) &= P\left(\frac{X-2}{5} > \frac{-2}{5}\right) = P\left(Z > \frac{-2}{5}\right) \\&= 1 - \Phi\left(-\frac{2}{5}\right) = \Phi\left(\frac{2}{5}\right)\end{aligned}$$

- (c) We have

$$\begin{aligned}P\left((X-2)^2 > 5\right) &= P\left(\frac{(X-2)^2}{25} > \frac{1}{5}\right) = P\left(Z^2 > \frac{1}{5}\right) \\&= P\left(Z > \frac{1}{\sqrt{5}}\right) + P\left(Z < -\frac{1}{\sqrt{5}}\right) \\&= 1 - \Phi\left(\frac{1}{\sqrt{5}}\right) + \Phi\left(-\frac{1}{\sqrt{5}}\right) \\&= 2\left(1 - \Phi\left(\frac{1}{\sqrt{5}}\right)\right)\end{aligned}$$

5.4 Example: Signal Transmission

- A binary message - either 0 or 1 - is transmitted through the atmosphere from A to B. The value 2 is sent when the message is 1 and the value -2 is sent when the message is 0. At the location B of the receiver, the message received is corrupted by some channel noise; that is if the signal $X = x$ has been transmitted then at the receiver we observe

$$R = x + N$$

where the noise is assumed to be a standard normal r.v.

- At the receiver, the following decoding scheme is used. If $R \geq 0.5$ then we conclude that 1 has been transmitted. If $R < 0.5$ then we conclude that 0 has been transmitted.
- What is the probability of decoding correctly the transmitted message when we transmit 0 and when we transmit 1?

5.4 Example: Signal Transmission

- If we transmit 0, then $R = -2 + N$ is a normal r.v. of mean -2 and variance 1 so

$$\begin{aligned}P(R < 0.5) &= P\left(\frac{R + 2}{1} < \frac{0.5 + 2}{1}\right) \\&= P(Z < 1.25) = \Phi(1.25) \approx 0.8944\end{aligned}$$

- If we transmit 1, then $R = 2 + N$ is a normal r.v. of mean 2 and variance 1 so

$$\begin{aligned}P(R > 0.5) &= P\left(\frac{R - 2}{1} > \frac{0.5 - 2}{1}\right) \\&= P(Z > -1.5) = \Phi(1.5) \approx 0.9332\end{aligned}$$

5.4 Normal Approximation to the Binomial Distribution

- Consider X a binomial r.v. of parameters n, p then we know that

$$E(X) = np, \text{Var}(X) = np(1-p).$$

- We have already seen that it is possible to approximate X by a Poisson distribution of parameter $\lambda = np$.
- As $n \rightarrow \infty$, it can be shown that X can be approximated by a normal r.v. with $\mu = np$ and $\sigma^2 = np(1-p)$ so

$$\begin{aligned} P(a \leq X \leq b) &= P\left(a \leq \frac{X - np}{\sqrt{np(1-p)}} \leq b\right) \\ &\approx \Phi(b) - \Phi(a). \end{aligned}$$

5.4 Example: Color blindness

- If 10% of men have color blindness, what is the probability that fewer than 100 in a random sample of 818 men are color blind?
- Let X be the number of color blindnessmen in a random sample of 818 men, this is a Bernoulli r.v. of parameters $p = 0.1$ and $n = 818$.
- We are interested in computing $P(X \leq 100)$. We can use the standard binomial but this is tedious. We use the normal approximation where

$$\mu = np = 81.8, \sigma = \sqrt{np(1-p)} = 8.5802$$

so

$$\begin{aligned} P(0 \leq X \leq 100) &= \Phi\left(\frac{100 - 81.8}{8.5802}\right) - \Phi\left(\frac{-81.8}{8.5802}\right) \\ &\approx 0.983. \end{aligned}$$

5.4 Example: Threshold signal

- Assume to transmit a random signal X which follows a normal distribution (μ, σ^2) . The receiver only detects signals above a given threshold m so that what is observed is

$$Y = \begin{cases} X & \text{if } X \geq m \\ 0 & \text{if } X < m \end{cases}$$

- Compute the expected value of the received signal Y ?

5.4 Example: Threshold signal

- We have

$$\begin{aligned} E(Y) &= \int_m^\infty x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= \int_m^\infty (x - \mu) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx + \mu \int_m^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \end{aligned}$$

where we use $x = (x - \mu) + \mu$.

- Now we have $\int_m^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1 - \Phi((m - \mu)/\sigma)$ and

$$\begin{aligned} &\int_m^\infty (x - \mu) \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= \left[\frac{-\sigma^2}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \right]_m^\infty \\ &= \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) \end{aligned}$$

so

$$E(Y) = \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) + \mu (1 - \Phi((m - \mu)/\sigma)).$$

5.5 Exponential distribution

- We have previously discussed the Poisson distribution. If the average rate of events is λ in a given time interval then it can be shown that the interarrival time between two events is a continuous r.v. with a so-called exponential density

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x \in [0, \infty) \\ 0 & \text{otherwise} \end{cases}.$$

- Example: Waiting times.

5.5 Example I

- Assume you receive an average of 2 phones calls per hour. You make the Poisson assumption so the time between two events follows an exponential density. Assume you have just hung up the phone,
 - (a) what is the probability that you are going to have to wait more than one hour to receive the next call?
 - (b) what is the probability that you are going to receive a phone call in the next ten minutes?
 - (c) what is the probability you are going to have to wait more than 10 minutes but less than one hour?

5.5 Example I

- Let X the time in hour you have to wait, this is a r.v. with an exponential pdf of parameter $\lambda = 2$.
- (a) We first want to compute $P(X > 1)$. We have

$$P(X > 1) = \lambda \int_1^{\infty} \exp(-\lambda x) dx = \exp(-\lambda) \approx 0.13.$$

- (b) We want to compute $P(X < \frac{10}{60})$ which is given by

$$P\left(X < \frac{10}{60}\right) = \lambda \int_0^{1/6} \exp(-\lambda x) dx = 1 - \exp(-\lambda/6) \approx 0.28.$$

- (c) We want to compute $P(\frac{10}{60} < X < 1)$

$$\begin{aligned} P\left(\frac{10}{60} < X < 1\right) &= P(X < 1) - P\left(X < \frac{10}{60}\right) \\ &= (1 - P(X > 1)) - P\left(X < \frac{10}{60}\right) \\ &\approx 0.58. \end{aligned}$$

5.5 Example II

- Assume a nuclear power plant has three independent safety systems. These safety systems have lifetimes X_1, X_2, X_3 in years which are exponential r.v.s with respective parameters $\lambda_1 = 1$, $\lambda_2 = 0.5$ and $\lambda_3 = 0.1$. Since their installation five years ago, these systems have never been inspected. What is the proba that the nuclear power plant is currently being operated without any working safety system?
- The probability that the safety system i it is not working is

$$\begin{aligned}\Pr(X_i < 5) &= \lambda_i \int_0^5 \exp(-\lambda_i x) dx = 1 - \exp(-\lambda_i 5) \\ &= \begin{cases} 0.9933 & \text{if } i = 1 \\ 0.9179 & \text{if } i = 2 \\ 0.3935 & \text{if } i = 3 \end{cases}\end{aligned}$$

- Hence the probability that none of the system is working is simply

$$\Pr(X_1 < 5) \Pr(X_2 < 5) \Pr(X_3 < 5) = 0.3588$$

5.5 Exponential distribution: hazard function

- Consider a positive continuous random variable X that we interpret as being the lifetime of some item. Let X have distribution function F and density f . The hazard rate (sometimes called the failure rate) function $\lambda(t)$ of F is defined by

$$\lambda(t) = f(t) / 1 - F(t).$$

- To interpret $\lambda(t)$, suppose that the item has survived for a time t and we desire the probability that it will not survive for an additional time dt . That is, consider

$$P\{X \in (t, t + dt) | X > t\}.$$

5.5 Exponential distribution: hazard function



$$\begin{aligned}P\{X \in (t, t + dt) | X > t\} &= \frac{P\{X \in (t, t + dt), X > t\}}{P\{X > t\}} \\&= \frac{P\{X \in (t, t + dt)\}}{P\{X > t\}} \\&\approx \frac{f(t)dt}{1 - F(t)}\end{aligned}$$

- Thus, $\lambda(t)$ represents the conditional probability intensity that a t -unit-old item will fail.
- For the exponential distribution with parameter λ , the hazard function is

$$\lambda(t) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$$

- The hazard (failure rate) function $\lambda(t)$ uniquely determines the distribution F . (Why?)

5.6 Other Continuous Distribution

- **Example: Log-normal Distribution.** X follows a log-normal(μ, σ^2) distribution if its pdf

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} \exp\left\{-\frac{1}{2\sigma^2}(\log x - \mu)^2\right\} & 0 < x < \infty, \\ 0 & \text{otherwise,} \end{cases}$$

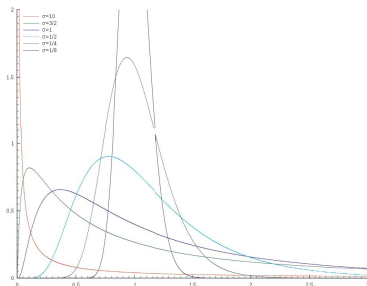


Figure: PDF of log-normal distribution.

5.6 Other Continuous Distributions

- If X follows a log-normal(μ, σ) distribution, then $\log X \sim N(\mu, \sigma^2)$.
- In finance, we often assume price P_t follows stochastic diffusion equation

$$dP_t = \mu P_t dt + \sigma P_t dW_t,$$

where W_t is a Brownian motion, μ and σ are constants. Using Ito's lemma,

$$d \log P_t = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t.$$

Here $\log P_t$ is normal distributed, and P_t is log-normal distributed.

5.6 Other Continuous Distributions

- **Example: Gamma Distribution.** A c.r.v. X follows a $\text{Gamma}(\alpha, \lambda)$ distribution if

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} & 0 < x < \infty, \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha, \beta > 0$, $\Gamma(\alpha)$ is the Gamma function $\int_0^\infty t^{\alpha-1} e^{-t} dt$.

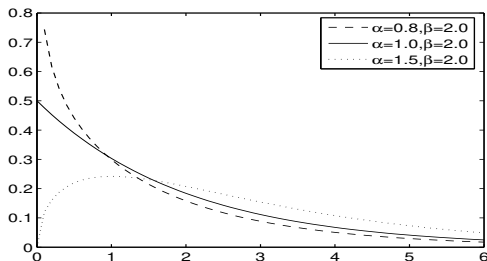


Figure: PDF of Gamma distribution.

5.6 Other Continuous Distributions

- $\text{Gamma}(1, \lambda) = \text{Exponential}(\lambda)$.
- **Properties of $\Gamma(\alpha)$**
 - (1) $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$;
 - (2) $\Gamma(k) = (k - 1)!$ where k is a positive integer;
 - (3) $\Gamma(1/2) = \sqrt{\pi}$.
- Two parameters— α : shape parameter; β : scale parameter.

5.6 Other Continuous Distributions

- **Example: Beta Distribution.** X follows a $\text{Beta}(\alpha, \beta)$ distribution if

$$f_X(x) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} & 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha, \beta > 0$, and $B(\alpha, \beta)$ is the Beta function

$$\int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx.$$

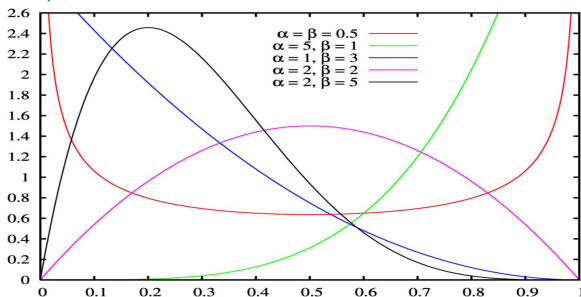


Figure: PDF of Beta distribution.

5.6 Other Continuous Distributions

- The support of *Beta* distribution is $[0,1]$.
- $Beta(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. (How to prove it?)
- $U(0, 1)$ is a special *Beta* distribution with $\alpha = \beta = 1$.

5.7 The Distribution of a Function of a Random Variable

- Let X be a r.v. of pdf $f_X(x)$ and consider the r.v. $Y = g(X)$.
- A legitimate question is to ask what is the pdf $f_Y(y)$ of Y .
- This has numerous applications: converting measurements, computing returns on investments etc.

5.7 Example: Conversion Celsius to Fahrenheit

- Consider X the temperature at a given time instant in Celsius. It is assumed that X has a pdf $f_X(x)$ and associated distribution function $F_X(x)$. Assume you want to convert this temperature in Fahrenheit, hence you introduce the r.v.

$$Y = \frac{9}{5}X + 32.$$

What is the distribution $F_Y(y)$ and its associated density $f_Y(y)$?

- Consider the general case where $Y = aX + b$ then

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr(aX + b \leq y) = \Pr\left(X \leq \frac{y-b}{a}\right) \\ &= F_X\left(\frac{y-b}{a}\right). \end{aligned}$$

- We know that $f_Y(y) = \frac{dF_Y(y)}{dy}$ so by the chain rule

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{a} f_X\left(\frac{y-b}{a}\right).$$

5.7 Example: Polar to Cartesian coordinates

- Assume you evolve on a circle of radius $R = 1$. The cartesian coordinates associated to an angle θ are

$$X = \cos \theta, \quad Y = \sin \theta.$$

Assuming that the angle θ is distributed uniformly on $[0, \frac{\pi}{2}]$, what is the pdf of Y ?

- We have $P(Y \leq y) = 0$ for $y \leq 0$ and $P(Y \leq y) = 1$ for $y \geq 1$.
For $0 \leq y \leq 1$

$$P(Y \leq y) = P(\sin \theta \leq y) = P(\theta \leq \sin^{-1} y) = \frac{2}{\pi} \sin^{-1} y.$$

- By differentiating $P(Y \leq y)$, we obtain

$$f_Y(y) = \frac{2}{\pi \sqrt{1-y^2}} \text{ for } 0 \leq y \leq 1$$

5.7 Example: Squared random variable

- Consider a r.v. X of pdf $f_X(x)$ and associated distribution function $F_X(x)$. You are not interested in X per se but in $Y = X^2$. What is the distribution $F_Y(y)$ and its associated density $f_Y(y)$?
- For $y \leq 0$, we have $F_Y(y) = 0$ and $f_Y(y) = 0$. For $y \geq 0$, we have

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr(X^2 \leq y) \\ &= \Pr(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

- Using the chain rule, we obtain

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}).$$

5.7 The Distribution of a Function of a Random Variable

- Consider the case where $g(x)$ is a strictly monotonic differentiable function of x . Then the r.v. $Y = g(X)$ has a pdf given by

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

- Proof follows for increasing $g(x)$ from

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \leq g^{-1}(y)) \text{ if } y = g(x) \text{ for some } x \\ &= F_X(g^{-1}(y)) \end{aligned}$$

- By the chain rule, we obtain

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dg^{-1}(y)}{dy} f_X(g^{-1}(y))$$

5.7 Example: Investment

- An investment account earns an annual interest rate R that follows a uniform distribution on the interval $(0.04, 0.08)$. The value of a 10,000\$ initial investment in this account after one year is given by $V = 10,000 \exp(R)$. Determine the cdf and pdf of V .
- We have $F_V(v) = 0$ for $v \leq 10,000 \exp(0.04)$ and for $v \geq 10,000 \exp(0.08)$

$$\begin{aligned} F_V(v) &= P(V \leq v) = P(R \leq \log v - \log 10,000) \\ &= \frac{1}{0.04} \int_{0.04}^{\log v - \log 10,000} dr = 25 \log(v) - 25 \log(10,000) - 1 \\ &= 25 \left[\log \frac{v}{10,000} - 0.04 \right] \end{aligned}$$

$$\text{so } f_V(v) = \frac{dF(v)}{dv} = \frac{25}{v}$$

- or directly for $g(r) = 10,000 \exp(r) = v \Leftrightarrow r = g^{-1}(v) = \log \frac{v}{10,000}$

$$f_V(v) = \left| \frac{dg^{-1}(v)}{dv} \right| f_R(g^{-1}(v)) = \frac{1}{v} \frac{1}{0.04} = \frac{25}{v}$$

5.7 Example: Cauchy distribution

- A real r.v. X follows Cauchy distribution if its pdf

$$f_X(x) = \frac{1}{\pi} \frac{1}{(1+x^2)}$$

What is the pdf of the r.v. $Y = g(X) = 1/X$?

- We have $X = g^{-1}(Y) = 1/Y$

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| \\ &= f_X(1/y) \left| -\frac{1}{y^2} \right| = \frac{1}{\pi} \frac{1}{(1+1/y^2)} \times \frac{1}{y^2} \\ &= \frac{1}{\pi} \frac{1}{(1+y^2)} = f_X(y). \end{aligned}$$