Chapter 4

STAT 5700: Probability

Prof. Katie Fitzgerald, PhD

Fall 2025

Contents

4.2 Continuous Random Variables

Discrete vs. Continuous

So far we have worked with **discrete** random variables that have probability distributions

$$\sum_S p(y) = 1$$

In this chapter we will consider **continuous** random variables that have **probability density functions** (pdfs)

$$\int_S f(y)dy = 1$$

Discrete random variables have a finite or countably infinite set of possible outcomes:

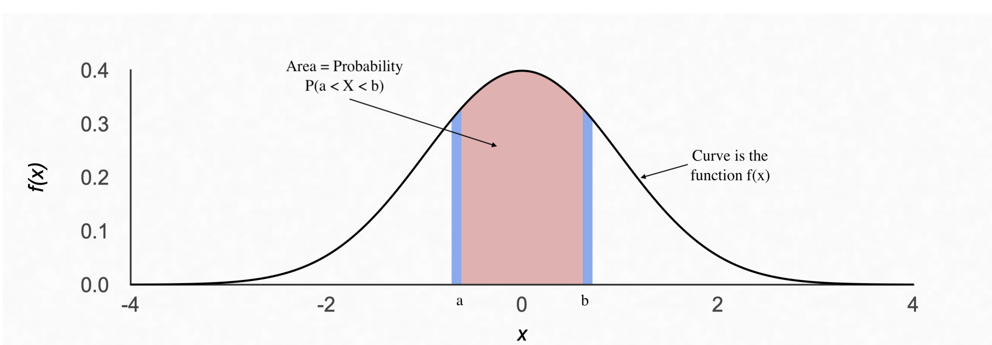
- Uniform (m-sided die): {1,2, . . . , m}
- Bernoulli (success/failure): {0,1}
- Binomial (# of successes in n trials): {0, 1, . . . n}
- Geometric (# of trials until 1st success): $\{1, 2, \dots\}$
- Negative Binomial (# of trials until rth success): $\{r, r + 1, r + 2, \dots\}$
- Hypergeometric (# of successes when sampling without replacement from finite population): {0, 1, 2, . . . r}
- Poisson (# of occurances in unit interval) $\{1, 2, \dots\}$

Continuous random variables, on the other hand, have an *interval* of possible outcomes, and decimal values are possible:

- amount of rain, in inches, that falls in a randomly selected storm
- weight, in pounds, of a randomly selected student
- square footage of a randomly selected house

The set of possible measurements in a continuous interval is (not countably) infinite and can't be put on a one-to-one correspondence with the integers.

> The probability that a continuous random variable Y takes on any particular value is 0, so we won't be finding $P(Y = y)$ for continuous random variables, but rather, $P(a < Y < b)$.

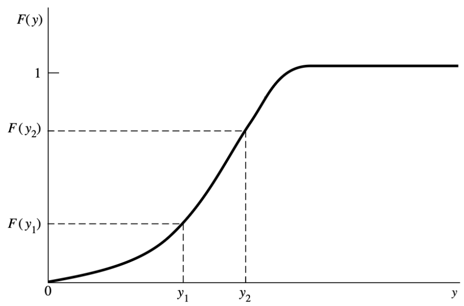


Recall: cumulative distribution functions (cdfs)

$$F(y) = P(Y \le y), \quad -\infty < y < \infty$$

is called the **cumulative distribution function (cdf)** of a random variable Y.

FIGURE **4.2**
Distribution function for a continuous random variable



Theorem 4.1: Properties of a cumulative distribution function: If $F(y)$ is a cdf, then

1. $F(-\infty) \equiv \lim_{y \to -\infty} F(y) = 0$.
2. $F(\infty) \equiv \lim_{y \to \infty} F(y) = 1$.
3. $F(y)$ is a nondecreasing function of y. If $y_1 < y_2$, then $F(y_1) \le F(y_2)$.

Properties of continuous pdfs and cdfs

Discrete	**Continuous**
p(y) = probability distribution	f(y) = probability density function (pdf)
$p(y) > 0, y \in S$	$f(y) \ge 0, y \in S$
$\sum_{y \in S} p(y) = 1$	$\int_S f(y)dy = 1$
If $A \subset S$, $P(Y \in A) = \sum_{Y \in A} p(y)$	If $(a, b) \subset S$, $P(a < Y < b) = P(a \le Y \le b) = \int_a^b f(y)dy$
cdf: $F(y) = P(Y \le y)$	cdf: $F(y) = P(Y \le y) = \int_{-\infty}^y f(t)dt, -\infty < y < \infty$

By the **Fundamental Theorem of Calculus**, we have that $F'(y) = f(y)$ for all y values where $F'(y)$ exists, and $P(a \le Y \le b) = F(b) - F(a)$

4.3 Expected value of continuous distributions

Expected value (including mean, variance, and moment generating functions) have the *same exact* definition for continuous random variables as they did for discrete random variables. Now, we're just evaluating them with **integrals** instead of **sums**.

Definition	**Discrete**	**Continuous**
$\mu = E(Y)$	$\sum_{y\in S} xp(y)$	$\int_S xf(y)dy$
$\sigma^2 = V(Y) = E[(Y-\mu)^2]$	$\sum_{Y\in S}(y-\mu)^2 p(y)$	$\int_S (y-\mu)^2 f(y)dy$
k^{th} moment $=E(Y^k)$	$\sum_{y\in S} y^k p(y)$	$\int_S y^k f(y)dy$
$m(t) = E(e^{tY})$	$\sum_{y\in S} e^{ty} p(y)$	$\int_S e^{ty} f(y)dy$
$E(g(Y))$	$\sum_{y\in S} g(y)p(y)$	$\int_S g(y)f(y)dy$

The following relationships still hold true for continuous random variables as well:

- standard deviation $= \sigma = \sqrt{\sigma^2} = \sqrt{V(Y)}$
- $V(Y) = E(Y^2) - [E(Y)]^2$
- first moment: $\mu = E(Y) = m'(0)$
- second moment: $E(Y^2) = m''(0)$
- $\sigma^2 = V(Y) = m''(0) - [m'(0)]^2$
- $E(cg(Y)) = cE(g(Y))$
- $E[c_1 g_1(Y) + c_2 g_2(Y)] = c_1 E(g_1(Y)) + c_2 E(g_2(Y))$

4.4 Uniform Distribution (continuous)

Example
Suppose it takes between 0 and 40 seconds for an elevator to arrive once you have pushed the button. We will assume that all wait times are equally likely. Define a random variable and determine a pdf to model this situation. *Hint: draw a picture.*

Example
Let the random variable Y denote the outcome when a point is selected at random from an interval $[a, b]$, $-\infty < a < b < \infty$. What might be a reasonable expression for the probability that the point is selected from the interval $[a, y]$, $\quad a \leq y \leq b$. *Hint: draw a picture.*

The *cdf* of a **continuous uniform random variable** Y is given by

$$F(y) = \begin{cases} 0, & y < a \\ \frac{y-a}{b-a}, & a \leq y \leq b, \\ 1, & b \leq y, \end{cases}$$

and the *pdf* of Y is given by

$$f(y) = \frac{1}{b-a}, \qquad a \leq y \leq b.$$

Similar to the discrete case, a random variable Y has a **uniform distribution** if its pdf is equal to a constant on its support.

However, in the continuous case, the support is a continuous interval rather than a discrete set of possible values. In the definition on the previous slide, the support of Y is the interval $[a, b]$.

We write $Y \sim U(a, b)$, to mean Y is a uniform random variable (or "is uniformly distributed") on the interval (a, b).

You will derive the mean, variance, and mgf of the continuous uniform distribution in the group work.

Elevator Example (cont'd)
Find and graph the cdf for the uniform random variable in the elevator example. What is the probability that the elevator arrives in less than 10 seconds? Between 15 and 20 seconds?

Exercise

Let Y have a uniform distribution $U(0,1)$ (we call this the "standard uniform distribution").

1. Write out the pdf and cdf of Y

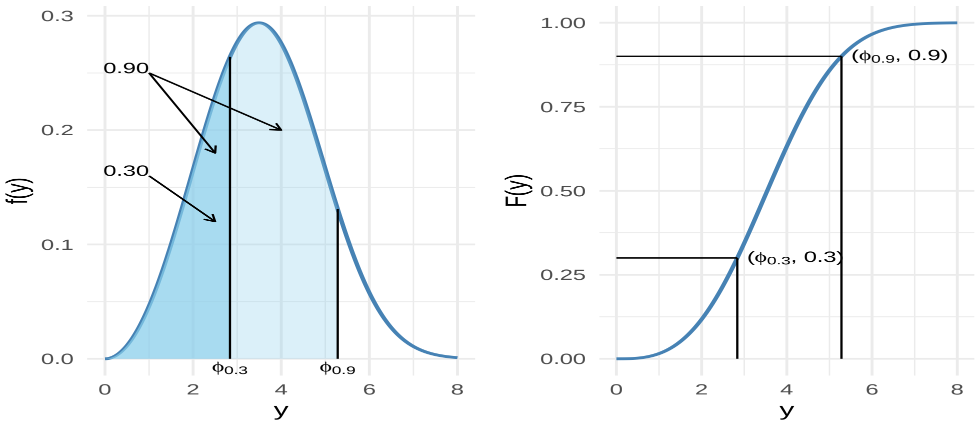
2. Let $W = a + (b-a)Y, \qquad a < b$ Find the cdf of W. How is W distributed?

Percentiles

The **(100p)th percentile** is a number ϕ_p, such that the area under $f(y)$ to the left of ϕ_p is p. That is,

$$F(\phi_p) = \int_{-\infty}^{\phi_p} f(y)dy = p$$

For example, the figure below illustrates the 30th ($\phi_{0.3}$) and 90th ($\phi_{0.9}$) percentiles.



The 50th percentile is called the **median**, $\phi_{0.50}$. The 25th and 75th percentiles are called the **first** and **third quartiles**, respectively, and are denoted by $q_1 = \phi_{0.25}$ and $q_3 = \phi_{0.75}$.

Exercise
The time Y in months until the failure of a certain product has pdf and cdf

$$f(y) = \frac{3y^2}{4^3} e^{-(y/4)^3}, \qquad 0 < y < \infty$$

$$F(y) = \begin{cases} 0, & -\infty < y < 0, \\ 1 - e^{-(y/4)^3} & 0 \le y < \infty \end{cases}$$

This is the distribution depicted in the figure above.
Find the median of this distribution.

4.6 The Exponential, Gamma, and Chi-square Distributions

A Poisson distribution is a discrete distribution used to model the number of occurrences in a given time interval. In particular, let Y be the number of occurences in an interval of length t, then Y is a Poisson random variable with pmf

$$f(y) = \frac{(\lambda t)^y e^{-\lambda t}}{y!}, \qquad y = 0, 1, 2, ...$$

λ is interpreted as the mean number of occurrences in the unit interval [0,1]. The number of occurrences is a discrete random variable, but the waiting times between successive occurrences are also random variables. However, these waiting times are continuous.

Exercise
Let W denote the waiting time until the first occurrence for a Poisson process in which the mean number of occurrences in the unit interval is λ.
What is the probability the time until first occurrence is less than w?

Use the above result to produce the pdf of W

Exponential Distribution

We just derived the pdf for what's known as the **exponential distribution**, which is used to model the waiting time for a first occurrence. We usually re-parameterize and let $\lambda = \frac{1}{\theta}$, because doing so gives the convenient interpretation of θ as the mean waiting time, $E(Y)$. We will derive this result soon.

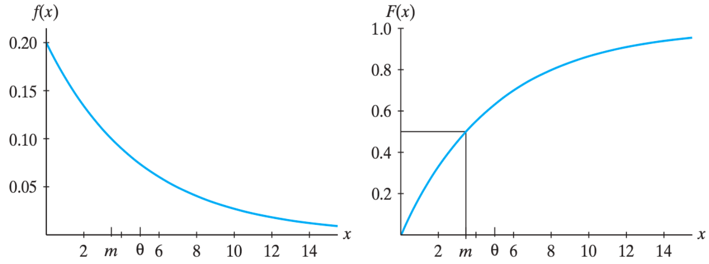


Figure 3.2-1 Exponential pdf, $f(x)$, and cdf, $F(x)$

Y has an **exponential distribution** if its pdf is defined by

$$f(y) = \frac{1}{\theta} e^{-y/\theta}, \qquad 0 \le y < \infty$$

and the cdf is defined by

$$F(y) = \begin{cases} 0 & y < 0 \\ 1 - e^{-y/\theta} & 0 \le y < \infty \end{cases}$$

We say $Y \sim exp(\theta)$.
Its mgf, mean, and variance are given by:

$$m(t) = \frac{1}{1 - \theta t}, \qquad t < \frac{1}{\theta}$$

$$E(Y) = \theta$$

$$V(Y) = \theta^2$$

Use the mgf to derive the mean and variance of the exponential distribution. That is, show that for $Y \sim exp(\theta)$, then $E(Y) = \theta$ and $V(Y) = \theta^2$. *See textbook for derivation of the mgf itself*

Example

Customers arrive in a certain shop according to an approximate Poisson process at a mean rate of 20 per hour (i.e. 1/3 per minute). What is the probability that the shopkeeper will have to wait more than 5 minutes for the arrival of the first customer? Also find the median time until the first arrival.

Gamma Distribution

The **Gamma distribution** is used to model the time until the αth occurrence in a Poisson process (with mean λ arrivals per unit.) The distribution has two parameters, α (the number of occurences) and $\theta = 1/\lambda$. We write $Y \sim gamma(\alpha, \theta)$ or sometimes $Y \sim \Gamma(\alpha, \theta)$ using the capital Greek letter Gamma (Γ).

$$f(y) = \frac{1}{\Gamma(\alpha)\theta^\alpha} y^{\alpha-1} e^{-y/\theta}, \qquad 0 \le y < \infty$$

$$m(t) = \frac{1}{(1-\theta t)^\alpha}, \qquad t < 1/\theta$$

$$E(Y) = \alpha\theta$$

$$V(Y) = \alpha\theta^2$$

The pdf of a gamma random variable involves what's known as the **gamma function**: $\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \qquad 0 < t$

When n is a positive integer, it can be shown that $\Gamma(n) = (n-1)!$. Typically, we will use R to evaluate the gamma function, particularly when n is not an integer.

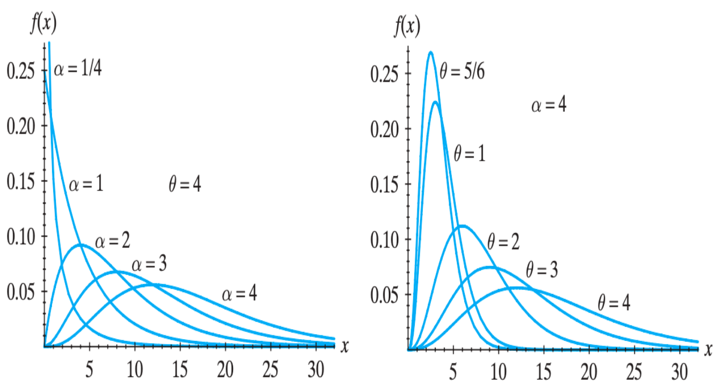


Figure 3.2-2 Gamma pdfs: $\theta = 4, \alpha = 1/4, 1, 2, 3, 4$; $\alpha = 4, \theta = 5/6, 1, 2, 3, 4$

Example

Suppose the number of customers per hour arriving at a shop follows a Poisson process with mean 30 people per hour. What is the distribution required to determine the probability the shopkeeper will wait more than 5 minutes before both of the first two customers arrive? What is the mean and variance of the random variable?

The Chi-square distribution

A special case of the **Gamma** distribution with $\alpha = \nu/2$ and $\theta = 2$ is known as the **Chi-square distribution**, where ν is a positive integer.
Let $Y \sim \Gamma(\nu/2, 2)$. The pdf of Y is

$$f(y) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}} y^{\nu/2-1} e^{-y/2}, \qquad 0 \leq y < \infty.$$

We say that Y has a **chi-square distribution with** ν degrees of freedom, which we abbreviate as $Y \sim \chi^2(\nu)$

Exercise

Using the properties of the gamma distribution, find the mean and variance of the chi-square distribution.

Continuous Random Variables in R

4.5 The Normal Distribution

The normal distribution has two parameters, μ and σ^2. We write $Y \sim N(\mu, \sigma^2)$. Y has a normal distribution if its pdf is defined by:

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

and the mgf is given by:

$$m(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

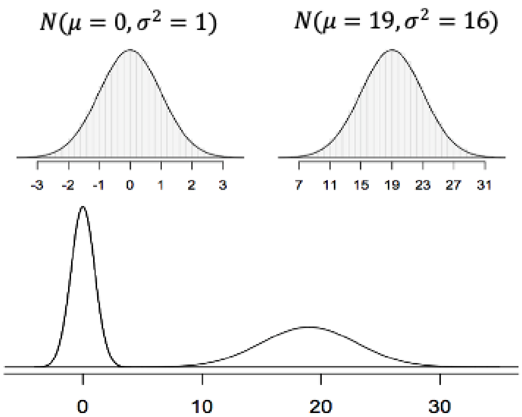
Normal mean and variance

Use the mgf to find the mean and the variance of the Normal distribution.

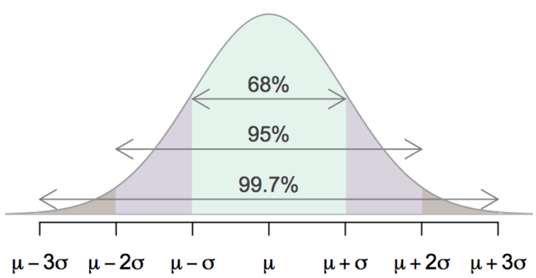
The normal distribution is unique in that its parameters in the pdf (μ and σ^2) are the same as its mean and variance.

This is NOT necessarily true for other distributions. For example, the Gamma distribution is defined by the two parameters α and θ, that is $Y \sim Gamma(\alpha, \theta)$, but $E(Y) = \alpha\theta \neq \alpha$ and $V(Y) = \alpha\theta^2 \neq \theta$

Properties of the Normal Distribution



- The normal distribution is unimodal and symmetric.
- The parameter μ controls the *center* of the distribution, and the parameter σ^2 controls the *spread* of the distribution.
- The normal distribution has a convenient property known as the **Emperical Rule**:
  - roughly 68% of the probability falls within 1 σ of the mean
  - roughly 95% falls within 2σ of the mean
  - roughly 99.7% falls within 3σ of the mean



Exercise:
SAT scores are distributed nearly normally with mean 1500 and standard deviation 300.
~68% of students score between ____________ and ____________ on the SAT.
~95% of students score between ____________ and ____________ on the SAT.
~99.7% of students score between ____________ and ____________ on the SAT

Standard Normal Distribution (Z)

The normal distribution with $\mu = 0$ and $\sigma^2 = 1$ is known as the **standard normal distribution**, and we usually use Z instead of Y to denote a standard normal random variable: $Z \sim N(0, 1)$.
The cdf of the standard normal distribution is also given special notation:

$$\Phi(z) = F(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\phi}} e^{-w^2/2} dw$$

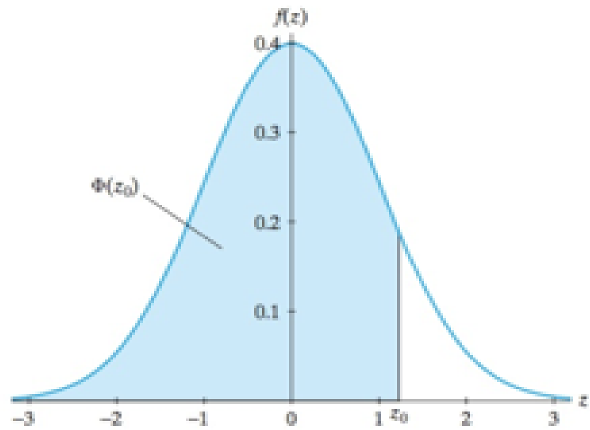


Figure 3.3-1 Standard normal pdf

We can convert any normal distribution into a standard normal distribution.

Let $Y \sim N(\mu, \sigma^2)$ and define the random variable $Z = \frac{Y-\mu}{\sigma}$

1. Show that $E(Z) = 0$

2. Show that $V(Z) = 1$

That is, the mean and the variance of Z are 0 and 1, respectively, no matter what the mean and variance of Y.

Z-scores

Z-scores are a way of understanding data as distances from the mean in "standard deviation units"

SAT scores are distributed nearly normally with mean 1500 and standard deviation 300. ACT scores are distributed nearly normally with mean 21 and standard deviation 5. A college admissions officer wants to determine which of the two applicants scored better on their standardized test with respect to the other test takers: Pam, who earned an 1800 on her SAT, or Jim, who scored a 24 on his ACT?

Properties of Z-scores

- Z-scores follow a standard normal distribution $Z \sim N(0, 1)$
- Z-scores will always be centered at ________________________
- Observations that fall BELOW the mean will have a ______________________ z-score
- Observations that fall ABOVE the mean will have a ______________________ z-score
- Observations that are more than 2 SD away from the mean (|Z| > 2) are usually considered unusual.
- Nearly all Z scores will be between _______________ and __________________ due to the ______________________________

Some important theorems

Theorem If $Y \sim N(\mu, \sigma^2)$, then $Z = (Y - \mu)/\sigma$ is $N(0, 1)$

Theorem If the random variable Y is $N(\mu, \sigma^2)$, $\sigma^2 > 0$, then the random variable $V = \frac{(Y-\mu)^2}{\sigma^2} = Z^2$ is a chi-square random variable with 1 degree of freedom. That is, $Z^2 \sim \chi^2(1)$

Chapter 4 Group Work

1. Find the mean, variance, and mgf for a continuous uniform distribution.

2. Let Y be a continuous random variable with pdf $f(y) = c(y - y^2), \quad 0 < y < 1$

   a. Find c
   b. Find $P(0.3 < Y < 0.6)$
   c. Find the median of Y

3. The pdf of Y is given by $f(y) = \frac{2}{y^3}, \qquad 1 < y < \infty$. Find $E(Y)$ and $V(Y)$.

4. For the function $f(y) = 4x^c, \;\; 0 < y < 1$,

   a. Find the constant c such that $f(y)$ is a valid pdf
   b. Find the cdf
   c. Find μ and σ
   d. Find $P(0.25 < Y < 0.75)$ using the cdf

5. Cars arrive at a tollbooth at a mean rate of five cars every ten minutes according to a Poisson process. Find the probability that the toll collector will have to wait longer than 26.30 minutes before collecting the 8th toll.