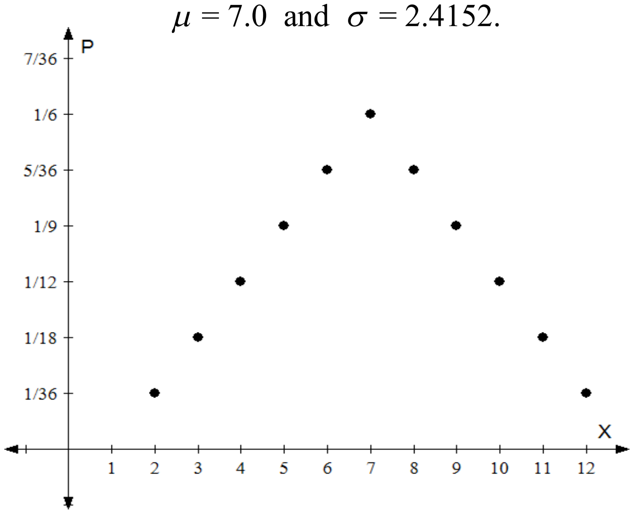
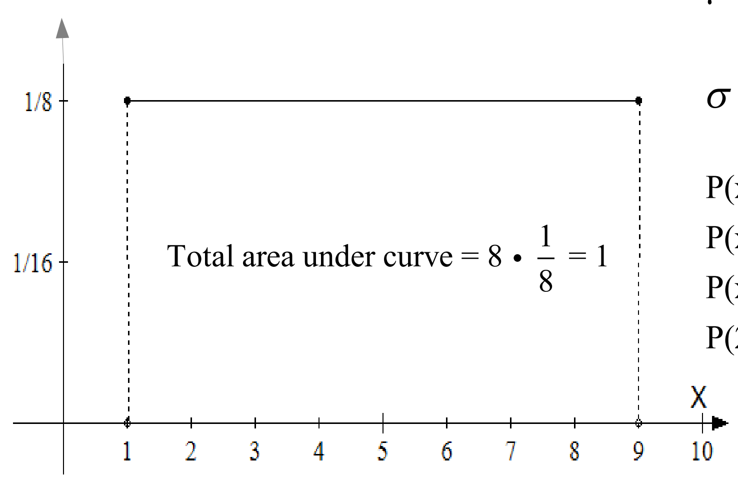
Summary of Common Probability Distributions

Description / Name	Probability Density Function	Expected Value	Standard Deviation	Comments
Toss a Single Fair Die x = the outcome of a toss. x = 1, 2, 3, 4, 5 or 6. P(x) = the probability the die toss landed **x**.	$P(x) = 1/6$	$\mu = 3.5$	$\sigma = \sqrt{\frac{35}{12}} \approx 1.7078$	The random variable **x** of the population of die tosses is discrete and the probability distribution of **x** is uniform.
Toss Two Fair Dice x = the outcome of a dice toss. x = 2, 3, 4, 5, . . . , 12. P(x) = the probability the total of the two dice = **x**.	$P(x) = \frac{6 - \lvert x - 7 \rvert}{36}$	$\mu = 7.0$	$\sigma = \sqrt{\frac{2 \cdot 35}{12}} \approx 2.4152$	The random variable **x** of the population of dice tosses is discrete and the probability distribution of **x** is symmetric. Refer to the graph of the distribution shown below. Page 8 of this handout shows how to derive μ and σ from μ and σ of a single die toss probability distribution.
Uniform Continuous Random variable **x** of the distribution is continuous and ranges from a to b.	$f(x) = \frac{1}{b - a}$ if $a \le x \le b$ $f(x) = 0$ if $x < a$ or $x > b$	$\mu = \frac{a+b}{2}$	$\sigma = \frac{b-a}{\sqrt{12}}$	Distribution is continuous. Probabilities of events are found by computing the area between the pd curve and the x-axis. The formula for μ makes intuitive sense, however, the formula for σ requires an understanding of integral calculus See the graph of the uniform continuous probability distribution below.

Graph of pd for tossing a pair of fair dice.
$\mu = 7.0$ and $\sigma = 2.4152$.



Graph of uniform continuous pd with
a = 1, b = 9, μ = 5.0 and $\sigma = 2.3094$.



$$\mu = \frac{1+9}{2} = 5$$

$$\sigma = \frac{9 - 1}{\sqrt{12}} = 2.3094$$

$P(x = 4) = 0$

$P(x < 4) = P(x \le 4) = 0.375$

$P(x > 7) = P(x \ge 7) = 0.25$

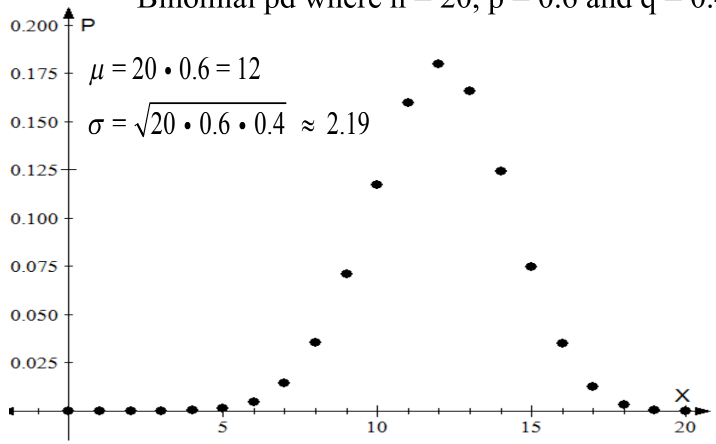
$P(2 \le x \le 8) = 0.75$

Description / Name	Probability Density Function	Expected Value	Standard Deviation
Binomial Prob. Distribution	$P(x) = {}_nC_x \cdot p^x \cdot q^{n-x}$	$\mu = np$	$\sigma = \sqrt{npq}$

A Bernoulli trial is an experiment that has two possible outcomes, named success or failure. The term 'success' does **not** necessarily mean that one of the outcomes is better than the other outcome; it is only a label that refers to one of the two possible outcomes.

n = the number of independent Bernoulli trials
p = the probability of success on each trial
q = the probability of failure on each trial.
x = the number of successes in **n** trials.
x = 0, 1, 2, 3, 4, . . . , **n**
P(x) = the probability of having exactly **x** successes in **n** Bernoulli trials.

Binomial pd where n = 20, p = 0.6 and q = 0.4



Binomial Distribution Requirements:

1) There must be a fixed number of trials.
2) Trials are independent and repeated under identical conditions.
3) Each trial has two possible outcomes, success or failure (S or F).
4) P(S) is the same for each trial. P(S) = p, P(F) = q and p + q = 1.

Example 1: The binomial pd would be a perfect model of the distribution of exam scores from a well constructed 25 question multiple choice exam on quantum mechanics with 5 choices per question. Each exam question is a Bernoulli trial. For most people, **p** = 0.2 and **q** = 0.8. For a university physics professor, **p** might equal 0.97 and **q** = 0.03. **P(x)** = the probability of correctly answering exactly **x** of the 25 questions.

Example 2: Consider a graduating class of 480 high school seniors and suppose the probability of a senior in the class reaching his/her 75th birthday equals 0.603. Each senior represents a Bernoulli trial and the number of Bernoulli trials equals 480. A senior that reaches his/her75th birthday represents a success and not reaching his/her 75th birthday represents a failure. **n** = 480, **p** = 0.603 . **q** = 0.397 and **P(x)** equals the probability that exactly **x** seniors in the class reached their 75th birthday.

Example 3: Toss 10 fair coins and let success equal a coin lands head. Each coin toss is a Bernoulli trial. **n** = 10 and **p** = **q** = 0.5. P(**x**) = the probability that **x** of the 10 coins landed head.

If np < 10 and n > 100, the Poisson pd with λ = np is a good approximation of the Binomial pd. Also if np > 10 and nq > 10, the normal pd with μ = np and σ = √(npq) is a good approximation of the binomial pd.

Poisson Distribution

$P(x) = \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}$	$\mu = \lambda$	$\sigma = \sqrt{\lambda}$

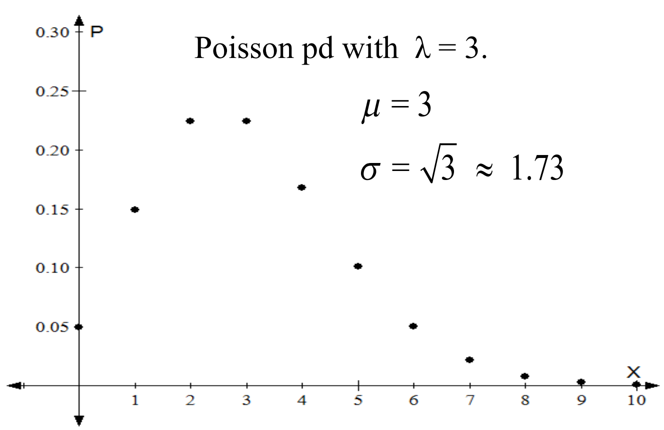
e is a famous math constant = 2.718281828459045 . . .

λ = the expected or average number of successes in an interval of length, area, volume, time, etc.

x = the number of successes found in the interval. The term 'success' refers to some outcome of interest.

x = 0, 1, 2, 3, 4, . . .

P(x) = the probability of finding exactly **x** successes in an interval of time, length, area, volume, etc.



The Poisson distribution answers questions about the probability of finding various numbers of successes in an interval of time, length, area, volume, etc. The following examples illustrate typical Poisson types of probability questions:

What is **P**(There are less than 3 defects in a 5,000 ft. coil of wire.) ?
What is **P**(Fred will catch 5 or more trout in Lake Lulu from 4:00 am to 8:00 am.)?
What is **P**(A container contains 4 to 10 deadly bacteria.)?
What is **P**(A page of the report contains at least one spelling error.)?
What is **P**(6 or less pieces of lost luggage in a group of 1,000 airline passengers.) ?

We assume that **λ** is <u>proportional</u> to the size of the interval and events are independent so that a success in an interval does not change the probability of another success in the same interval. For most problems, you must carefully read the problem and then set up and solve a proportion in order to find the correct value for the **λ** parameter.

Under the right conditions, the Poisson distribution is a good approximation of the Binomial distribution. Refer to the comment section of the Binomial distribution. Also if **λ** > 10, the normal pd with μ = **λ** and σ = √(**λ**) is a good approximation of the Poisson pd.

Description / Name	Probability Density Function	Expected Value	Standard Deviation
Geometric Probability	$P(x) = q^{x-1} \bullet p$	$\mu = \frac{1}{p}$	$\sigma = \frac{\sqrt{1 - p}}{p} = \frac{\sqrt{q}}{p}$

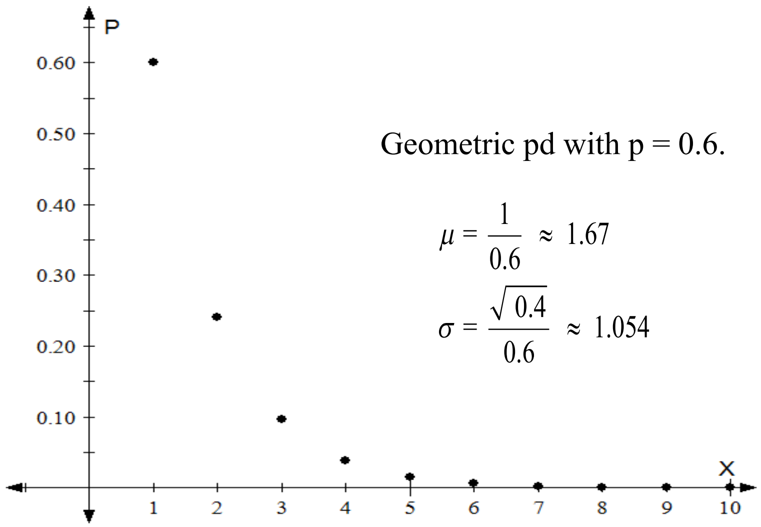
p = probability of success on each trial. The term 'success' refers to an outcome of interest.

q = probability of failure on each trial. The term 'failure' refers to one of the two possible outcomes of interest on a trial.

x = the number of trials needed to achieve the first success.

x = 1, 2, 3, 4, . . .

P(x) = the probability of failure on the first **x-1** trials and a success on trial number **x**.



Geometric pd with p = 0.6.

$$\mu = \frac{1}{0.6} \approx 1.67$$

$$\sigma = \frac{\sqrt{0.4}}{0.6} \approx 1.054$$

Properties and Requirements of the Geometric Distribution

a) Distribution is skewed to the right.
b) There is no fixed number of trials since trials are repeated until the first success is achieved. The minimum number of trials = 1.
c) Trials are independent and repeated under identical conditions.
d) Each trial has two possible outcomes, success or failure. (S or F)
e) P(S) is the same for each trial. P(S) = p, P(F) = q and p + q = 1.
f) x - 1 in the formula for the Geometric pdf represents the number of failures before the first success is achieved on trial number **x**.

The Geometric probability distribution is used to answer questions about the probability needing some number of attempts in order to achieve the first success. The following examples illustrate typical types of questions that can be answered with the Geometric pd:

a) What is **P**(It will take Fred 3 attempts to pass the CPA exam.) ?
b) What is **P**(It will take 3 or less tries for Linda to pass MTH 220)?
c) What is **P**(It will take Sue 6 or more phone calls to make her first cold call sale of the day)?

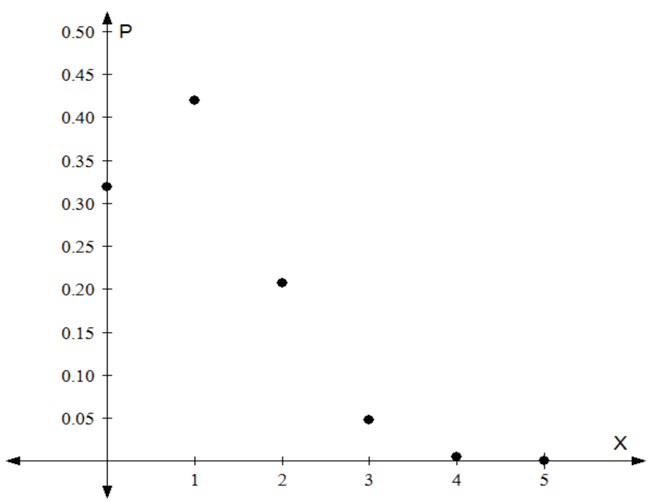
Hypergeometric pdf

$$P(x) = \frac{{}_aC_x \cdot {}_bC_{n-x}}{{}_NC_n} \qquad \mu = \frac{na}{N} \qquad \sigma = \sqrt{\frac{nab(N-n)}{N^2(N-1)}}$$

N = the population size
n = the random sample size
a = the total number of successes in the population (See below)
b = the total number of failures in the population (See below)
x = the number of successes found in the random sample of size **n** where the **s**ample is taken without replacement.
x = 0, 1, 3, . . . , j where $j \leq a$ and $j \leq n$.

P(x) = the probability of finding exactly **x** successes in a random sample of size **n**.

The terms 'success' and 'failure' refer to two properties of the population that are of interest.



Sample application: A lot of 100 fuses will be inspected by testing a random sample of 5 fuses. If all 5 "blow" at the correct amperage, the lot will be accepted. Suppose the lot contains 20 defective fuses and we want to find the probability of finding various numbers of defective fuses in a random sample of 5 fuses when sampling is done without replacement.

Since we a looking for defective fuses, 'success' equals a selected fuse is defective and 'failure' equals a selected fuse is not defective. See the graph to the left.

N = 100, n = 5, a = 20, b = 80, and x = 0, 1, 2, 3, 4 or 5. $P(x) = \frac{{}_{20}C_x \cdot {}_{80}C_{5-x}}{{}_{100}C_5}$

$$\mu = \frac{5*20}{100} = 1, \quad \sigma = \sqrt{\frac{5*20*80*(100-5)}{100^2(100-1)}} = 0.8762$$

$P(0) = 0.319309$, $P(1) = 0.420144$, $P(2) = 0.207344$

$P(3) = 0.047849$, $P(4) = 0.005148$, $P(5) = 0.000206$

Description / Name	Probability Density Function	Expected Value	Standard Deviation
Normal or Gaussian pdf	$P(x) = \dfrac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$	μ	σ

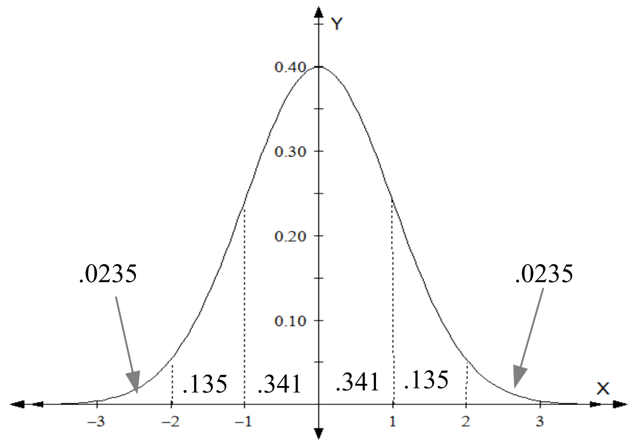
The random variable **x** of a normal distribution is a continuous and ranges from $-\infty$ to ∞.

The mean equals the μ parameter in the equation of the probability function.

The standard deviation equals the σ parameter in the equation of the probability function.

The **z-score** of **x** equals the number of standard deviations **σ** that **x** is above or below the mean μ. The properties of the NPD make the z-score the common currency of two different normal probability distributions.

Standard Normal Distribution (μ = 0 and σ = 1)



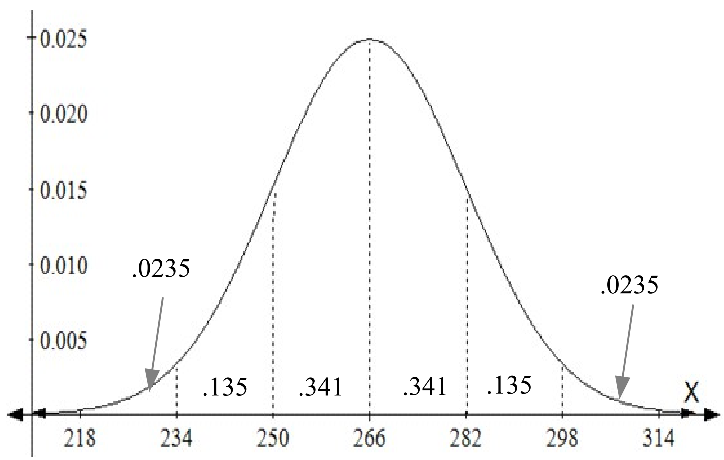
Standard normal horizontal scale values are also z-scores.

Z-score Conversion

$$z = \frac{x-\mu}{\sigma}$$

$$x = \mu + z\sigma$$

Normal Distribution with μ = 266 and σ = 16



Lengths of Pregnancies in Days for Humans

Comments

Background: The distribution of data values for a wide variety of populations can be described by the normal probability distribution. Consequently, many of the populations dealt with in tstatistics courses have a normal probability distribution or almost normal probability distribution. The great German mathematician Carl Friedrich Gauss (1777-1855) used the normal probability distribution to explain apparent errors in observations of star positions. Gauss explained that a star's position varies in an interval of space and therefore a star's position in space is not an absolute or fixed location . He showed that the normal probability distribution can be used to calculate the probability of observing a star in a particular region of space. Gaussian curve and the bell-curve are other names that refer to the normal probability distribution.

Properties of the Normal Probability Distribution

Every NPD has two parameters that completely describe the distribution. The mean μ is the center of the distribution and the standard deviation σ describes the variation of the distribution.

Every NPD is symmetrical about the mean μ.

Every NPD curve ranges from minus infinity to plus infinity and is above the horizontal axis. The graph never touches the horizontal axis even though it appears to touch the horizontal axis in most graphs of the NPD.

The total area under every NPD curve equals 1. (If you think about it, this is an incredible fact!)

The probability that a data value lies in an interval of finite or infinite length equals the area under the curve and above the interval.

For every NPD and every constant **k**, the probability that a data value lies within **k** standard deviations from the mean **μ** is always the same.

The NPD can be used to approximate a binomial probability distribution if $np \geq 10$ and $nq \geq 10$. Set $\mu = np$ and $\sigma = \sqrt{(npq)}$. Example: When n = 50 and p = 0.6, then μ = 30, and σ = 3.464. Using a continuity correction of 0.5 and a TI-graphing calculator we have:

a) $P(x \leq 17)$ = binomcdf (50,0.6,17) ≈ normalcdf (-0.5, 17.5, 30, 3.464)
b) $P(x \geq 8)$ = 1 - binomcdf (50,0.6,7) ≈ normalcdf (7.5, 50.5, 30, 3.464)
c) $P(8 \leq x \leq 28)$ = binomcdf (50, 0.6, 28) – binomcdf (50, 0.6, 7)
≈ normalcdf (7.5, 28.5, 30, 3.464).

Comment: Some authors only require $np \geq 5$ and $nq \geq 5$ If p = 0.9999 and n = 5,000, then np = 4,999.5 and nq = 0.5. In this case, the normal pd would be a very poor approximation of the binomial pd.

Description / Name | **Probability Density Function** | **Expected Value** | **Standard Deviation** | **Comments**

Student's t Distribution or **t -Distribution**

The random variable **t** of a **t**-distribution is continuous and ranges from $-\infty$ to ∞ .

The probability function for the **t**-distribution depends on a degrees of freedom parameter and the gamma function **Γ**(x). The ν parameter in the formula for f(t) is a positive integer which equals the degrees of freedom of the t-distribution.

Properties of Gamma Function
(Non calculus students can ignore.)

$$\Gamma(n) = \int_0^{\infty} t^{n-1} e^{-t} dt \quad n > 0$$

$\Gamma(1) = 1$
$\Gamma(n + 1) = n\Gamma(n)$
$\Gamma(n+1) = n!$ if n is a positive integer
$\Gamma(1/2) = \sqrt{(\pi)}$
$\Gamma(p)\Gamma(1-p) = \pi / \sin(\pi p)$

$$\Gamma\left(\frac{15}{2}\right) = \frac{13}{2}\bullet\frac{11}{2}\bullet\frac{9}{2}\bullet\frac{7}{2}\bullet\frac{5}{2}\bullet\frac{3}{2}\bullet\frac{1}{2}\sqrt{\pi}$$

$$\Gamma\left(\frac{16}{2}\right) = \Gamma(8) = 7! = 5{,}040$$

$$f(t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{t^2}{v}\right)^{-(v+1)/2}$$

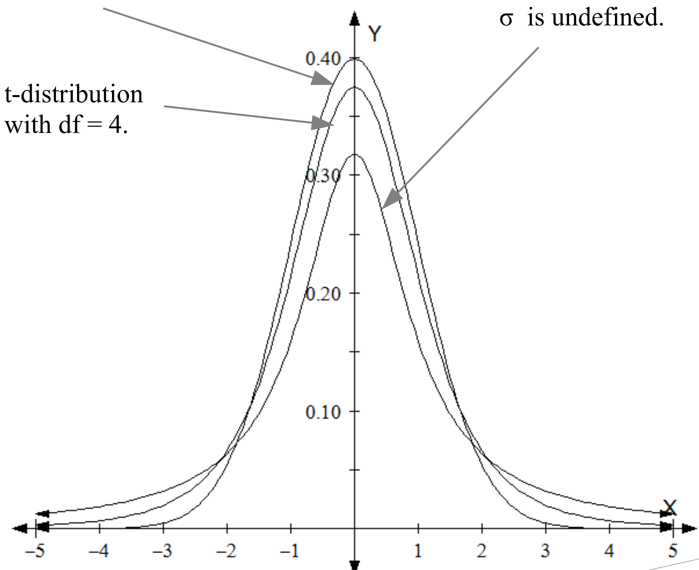
$$\mu = 0 \qquad \sigma = \sqrt{\frac{\nu}{\nu - 2}} \text{ if } \nu > 2$$

$$= \sqrt{\frac{df}{df - 2}} \text{ if } df > 2$$

Standard Normal Distribution
(**μ** = 0 and **σ** = 1)

t-distribution with df = 4.

t-distribution with df = 1. When df $\leq$ 2, σ is undefined.



Other statistics besides the t-statistic above have a **t** probability distribution.

Background: In 1908, W.S. Gosset (1876 – 1937) discovered an important sampling statistic **t** and its corresponding probability distribution which are useful in estimating a population μ without knowing the population σ. Gosset was employed as a statistician by the Guinness Brewing Company in Dublin, Ireland. Because his employer frowned on publication of research by its employees, Gossett published his results under the pen name 'Student'. The **t**-distribution is useful when the population σ is unknown or using small samples from normal or almost normal populations.

Properties of the t Distribution (n = the sample size. n should be greater than 30 or the parent population should be normal or almost normal.)

a) Mean **μ** = 0

b) Degrees of freedom parameter **df** of a t- distribution = **n** – 1 where **n** equals the sample size.

c) As the sample size **n** increases, the corresponding **t**-distribution approaches a standard normal pd.

d) The standard deviation **σ** is greater than 1. As the sample size increases, **σ** decreases and converges to 1.

e) The sampling statistic **t** below has a **t**-distribution with **n** – 1 degrees of freedom. The t-statistic is used to calculate confidence intervals for population means when σ is unknown and is the test statistic for several types of hypothesis tests. The formula for sampling statistic **t** is:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{x} - \mu)}{s}$$

$\bar{x}$ = the mean of a random sample
s = the standard deviation of a random sample
μ = the mean of the parent population
n = the size of the random sample

Description / Name | **Probability Density Function** | **Expected Value** | **Standard Deviation** | **Typical Applications**

Chi-Square Distribution

The random variable **x** of a **χ^2 distribution** is a continuous and ranges from 0 to ∞.

The probability function for the **chi-square, χ^2,** distribution depends on the gamma function, **Γ**(x), and a parameter **df** that is a positive integer which represents the degrees of freedom of the probability distribution. **df** equals the **ν** parameter in the formula for f(x).

Properties of Gamma Function

(Non calculus students can ignore these equations.)

$$\Gamma(n) = \int_0^{\infty} t^{n-1} e^{-t}\, dt \quad n > 0$$

$\Gamma(1) = 1$
$\Gamma(n + 1) = n\Gamma(n)$
$\Gamma(n+1) = n!$ if n is a positive integer
$\Gamma(1/2) = \sqrt{\pi}$
$\Gamma(p)\Gamma(1-p) = \pi / \sin(\pi p)$

$$\Gamma\left(\frac{15}{2}\right) = \frac{13}{2}\cdot\frac{11}{2}\cdot\frac{9}{2}\cdot\frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi}$$

$$\Gamma\left(\frac{16}{2}\right) = \Gamma(8) = 7! = 5{,}040$$

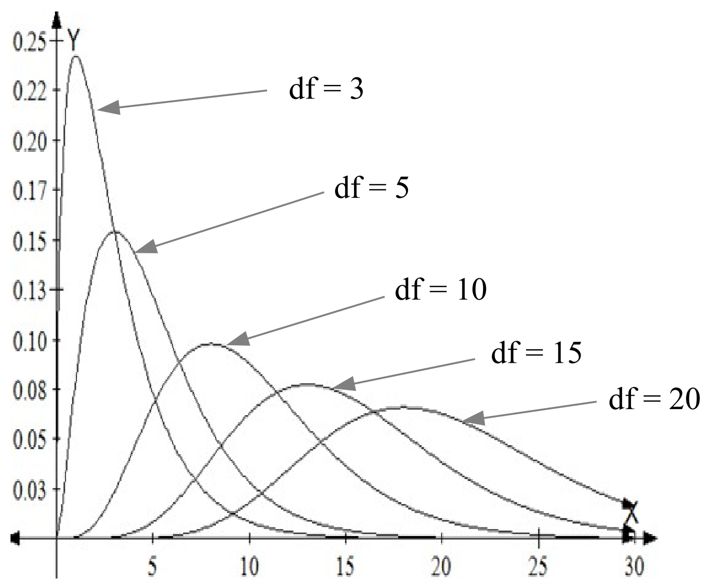
$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2 - 1} e^{-x/2} \quad if \;\; x \geq 0$$

$f(x) = 0 \quad$ if $x < 0$

$$\mu = v = df = n - 1$$

$$\sigma = \sqrt{2v} = \sqrt{2(df)} = \sqrt{2(n - 1)}$$

Chi-square distributions with 3, 5, 10, 15 and 20 degrees of freedom.



It so happens that many other sampling statistics have a χ^2 probability distribution.

Typical Applications

a) Compute confidence intervals for σ and σ^2 of a normal probability distribution and test claims about σ and σ^2.

b) Goodness-of-Fit: Determine if a an observed set of data values fit a specific probability distribution.

c) Determine whether or not two random variables are independent. Are events associated with one probability distribution independent of events associated with another probability distribution?

d) Test a claim involving homogeneity. Do difference populations have the same proportions of different characteristics?

Properties of the Chi-Square Distribution (n = sample size)

a) The number of degrees of freedom df = n – 1 and **μ** = df = n-1.

b) The chi-square distribution is not symmetric. As the number of degrees of freedom increases, the chi-square distribution approaches a normal pd.

c) If the pd of the parent population is normal, then the χ^2 sample statistic defined below has a chi-square pd with n-1 degrees of freedom.

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2} \quad \text{where } s^2 = \text{ sample variance}$$

and σ^2 = parent population variance

d) If all n independent random variables $\mathbf{x_i}$ have a standard normal pd (μ= 0 and σ = 1), then the χ^2 statistic defined below has a chi-square pd with n -1 degrees of freedom..

$$\chi^2 = x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2$$

Description / Name	Probability Density Function	Expected Value	Standard Deviation	Typical Applications

F-Probability Distribution

The random variable **x** of a F-distribution is a continuous and ranges from 0 to ∞ .

The probability function for the **F** distribution depends on the gamma function **Γ**(x), and two parameters **df1 and df2** that are positive integers which represent degrees of freedom of the probability distribution. The parameters ν_1 and ν_2 in the formula for f(x) equal df1 and df2 respectively.

$$f(x) = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \nu_1^{\nu_1/2} \nu_2^{\nu_2/2} x^{\nu_1/2 - 1} (x\nu_1 + \nu_2)^{-(\nu_1 + \nu_2)/2} \text{ if } x > 0$$

$$f(x) = 0 \quad \text{if } x \le 0$$

$$\mu = \frac{\nu_2}{\nu_2 - 2} \quad \text{for } \nu_2 > 2$$

$$\sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2} \quad \text{for } \nu_2 > 4$$

$$x_{mode} = \left(\frac{\nu_1 - 2}{\nu_1}\right)\left(\frac{\nu_2}{\nu_2 + 2}\right) \quad \text{for } \nu_1 > 2$$

Properties of Gamma Function

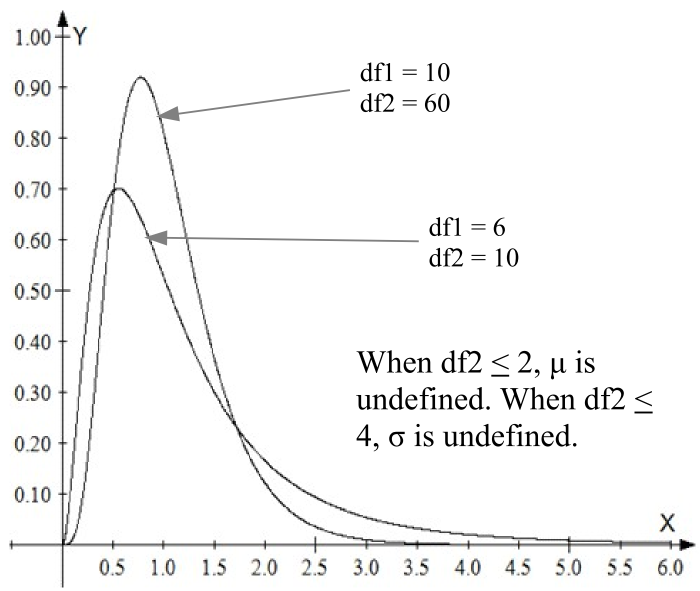
(Non calculus students can ignore these equations.)

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt \quad n > 0$$

Γ(1) = 1
Γ(n + 1) = nΓ(n)
Γ(n+1) = n! if n is a positive integer
Γ(1/2) = √(π)
Γ(p)Γ(1-p) = π / sin(πp)

$$\Gamma\left(\frac{15}{2}\right) = \frac{13}{2}\bullet\frac{11}{2}\bullet\frac{9}{2}\bullet\frac{7}{2}\bullet\frac{5}{2}\bullet\frac{3}{2}\bullet\frac{1}{2}\sqrt{\pi}$$

$$\Gamma\left(\frac{16}{2}\right) = \Gamma(8) = 7! = 5{,}040$$



When df2 ≤ 2, μ is undefined. When df2 ≤ 4, σ is undefined.

Typical Applications

a) Compare variances of two normal populations by testing claims about population variances. Let n_1 and n_2 sizes of random samples taken from two normal populations. df1 = n_1 - 1 and df2 = n_2 - 1. If s_1^2 and s_2^2 equal the sample variances, the test statistic F = $\frac{s_1^2}{s_2^2}$ has a F-distribution with df1 and df2 degrees of freedom.

b) One-Way ANOVA to determine whether or not the means of three or more normal treatment populations are equal. Let k = the number of treatment populations and n_1 n_2 . . . n_k sizes of the random samples taken from the k populations. Let N = n_1 + n_2 + . . . + n_k. df1 = k - 1 and df2 = N - 1. A complicated test statistic derived from the k samples has a F-distribution with df1 and df2 degrees of freedom.

c) Two-Way ANOVA to test for interaction between two factors such as two prescription drugs or gender of student and college major.

Basics of Linear Transformations of Data and Random Variables

Linear Transformation of a Data Set

Let s and s^2 equal the standard deviation and variance of a sample from a population.

If $x_1, x_2, x_3, \ldots, x_n$ is a sample of data values from a population and $y_i = a + bx_i$ where a and b are constants, then $\overline{y} = a + b\overline{x}$, $s_y^2 = b^2 s_x^2$ and $s_y = |b| s_x$.

Consider a sample of 12 adult male heights measured in inches: **68 69 71 70 73 67 70 74 65 67 65 69**

$\overline{x} = 69$ inches, s = 2.8284 inches and $s^2 = 8.0000$

If we increase each height in the sample by 4 inches, how would the sample mean and sample variation change?

$\overline{x} = 69$ inches, s = 2.8284 inches and $s^2 = 8.0000$

$y_i = 4 + 1 \bullet x_i = x_i + 4$

$\overline{y} = \overline{x} + 4 = 73$ inches

$s_y = 1 \bullet s_x = s_x = 2.82843$

The mean changes 4” and no change in variation.

If we increase each height in the sample by 4 inches and convert all heights to cm, how would the sample mean and sample variation change?

1 inch = 2.54 cm and 4 inches = 10.16 cm

$y_i = 10.16 + 2.54 \bullet x_i$

$\overline{y} = 2.54\overline{x} + 10.16 = 185.42$

$s_y = 2.54 \bullet s_x = 7.18421$

Linear Combination of n Independent Random Variables

Let $x_1, x_2, x_3, \ldots, x_n$ represent n independent random variables. The x_i's could be random variables from n different independent populations or obtained by sampling with replacement from a single population. When sampling is done without replacement from a single population, the x_i's can be treated as n independent random variables if the sample size is no more than 5% (some books use 10%) of the population size.

Let $\mu_1, \mu_2, \mu_3, \ldots, \mu_n$ equal the means of the n random variables.

Let $\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_n^2$ equal the variances of the n random variables.

Let $c_1, c_2, c_3, \ldots, c_n$ equal n real number constants.

General Case: Let the random variable $X = c_1x_1 + c_2x_2 + c_3x_3 + \ldots + c_nx_n$

$$\mu_X = c_1\mu_1 + c_2\mu_2 + c_3\mu_3 + \ldots + c_n\mu_n$$

$$\sigma_X^2 = c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + c_3^2\sigma_3^2 + \ldots + c_n^2\sigma_n^2$$

$$\sigma_X = \sqrt{c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + c_3^2\sigma_3^2 + \ldots + c_n^2\sigma_n^2}$$

The random variable X is a linear combination of **n** independent random variables.

Let the x_i's equal the values of a random sample of size n taken with replacement from the same parent population with mean = μ and population standard deviation = σ.

$$\text{If } X = \overline{x} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \frac{1}{n}x_3 + \ldots + \frac{1}{n}x_n = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}\text{, then}$$

$$\mu_{\overline{x}} = \frac{1}{n}\mu + \frac{1}{n}\mu + \frac{1}{n}\mu + \ldots + \frac{1}{n}\mu = \frac{n\mu}{n} = \mu$$

$$\sigma_{\overline{x}}^2 = \frac{1}{n^2}\sigma^2 + \frac{1}{n^2}\sigma^2 + \frac{1}{n^2}\sigma^2 + \ldots + \frac{1}{n^2}\sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Description / Name	Probability Density Function	Expected Value	Standard Deviation	Comments

Weibull Probability Distribution

The random variable **x** of a Weibull distribution is continuous and ranges from 0 to ∞ .

The properties of the distribution are completely determined by two positive real number constants, α and β.

If α < 1, the y-axis is a vertical asymptote of the distribution and the distribution is strictly decreasing for all x > 0.

If α = 1, the distribution is exponential with μ = σ = β and median = βln(2).

If α > 1, the distribution has a relative maximum when x = $\beta((\alpha - 1)/\alpha)^{1/\alpha}$.

μ, σ and the median of the distribution are all proportional to β.

α is called the shape parameter and β is called the location parameter.

$$\text{f(x)} = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^{\alpha}} \quad \text{if } x \ge 0$$

$$= 0 \quad \text{if } x < 0$$

$$\mu = \beta\cdot\Gamma\left(1+\frac{1}{\alpha}\right) \quad \text{where } \Gamma \text{ is the gamma function}$$

$$\sigma = \sqrt{\beta^2\left\{\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma\left(1+\frac{1}{\alpha}\right)^2\right\}}$$

$$\text{Median} = \beta\ln(2)^{1/\alpha}$$

$$\text{Mode} = \beta\left(\frac{\alpha-1}{\alpha}\right)^{1/\alpha} \quad \text{if } \alpha > 1$$

$$= 0 \quad \text{if } \alpha = 1$$

Some Properties of the Gamma Function

$$\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-x}dt \quad \text{for } x > 0$$

$$\Gamma(1) = 1$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$\Gamma(x+1) = x\Gamma(x) \quad \text{if } x > 0$$

$$\Gamma(x+1) = x! \quad \text{if } x \text{ is a positive integer}$$

$$\Gamma(p)\Gamma(1-p) = \pi / \text{Sin}(\pi p)$$

The Weibull distribution is very useful because it can be made to fit a wide variety of data sets by varying the values of α and β.

Common applications involve modeling the lifetime of components such as bearings, ceramics, capacitors, and dielectrics.

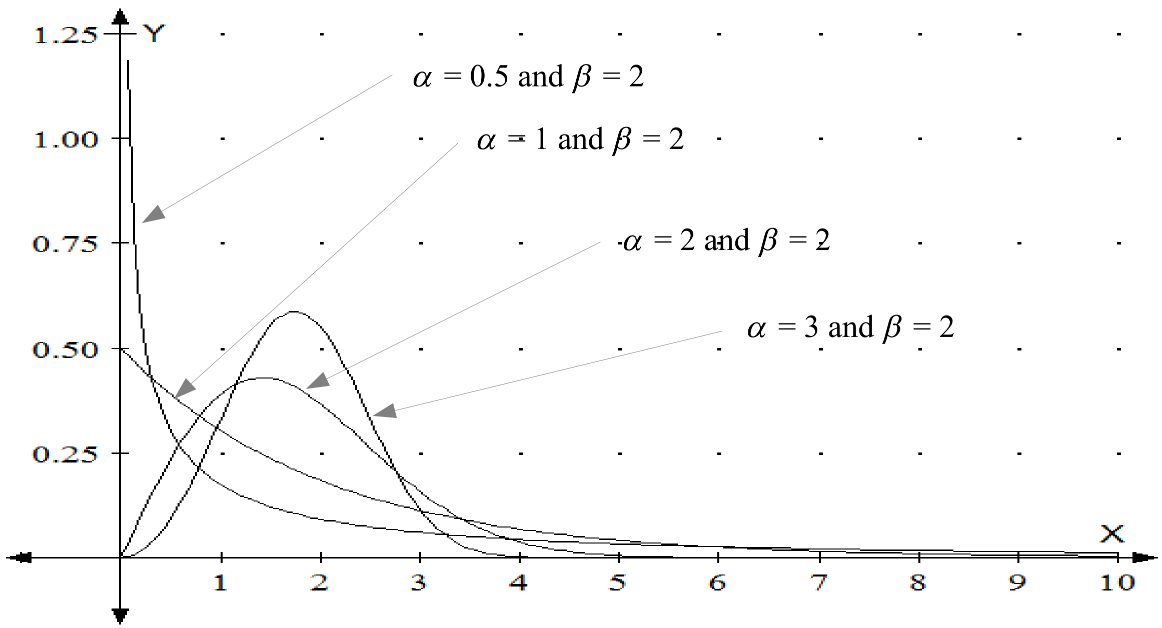
Other uses of the Weibull distribution include industrial engineering to study manufacturing and delivery times, weather forecasting, insurance industry to model the size of various types of insurance claims, hydrology, and the study of the size of particles resulting from milling, grinding and crushing operations.

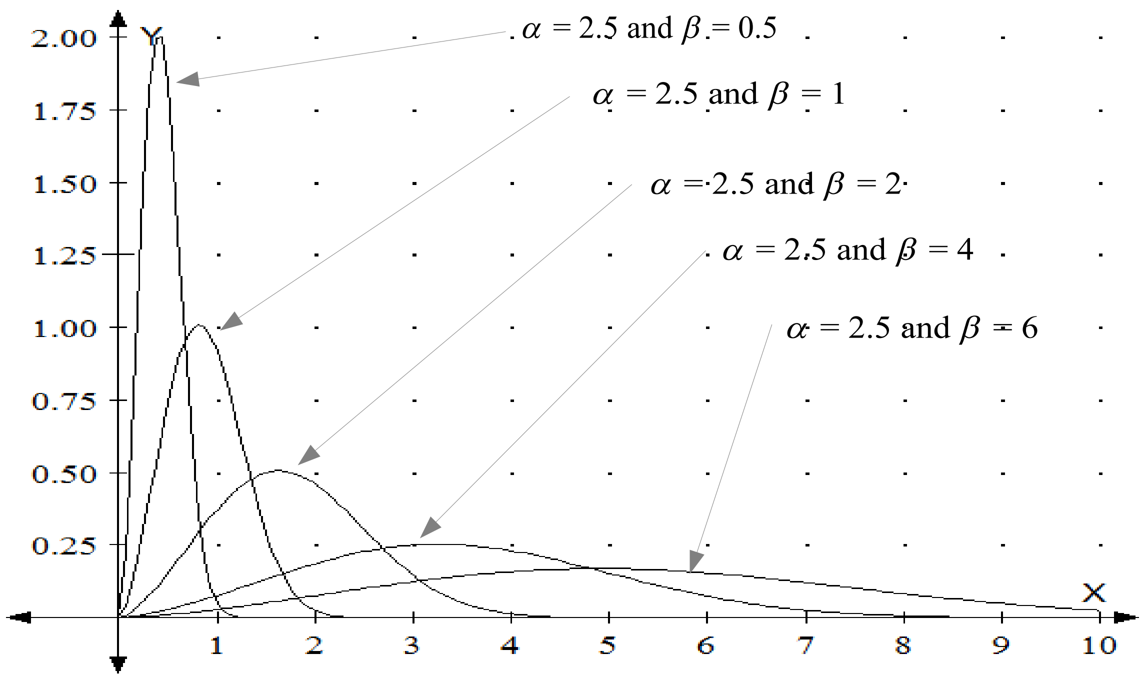
If μ and σ are known, the relationship

$$\frac{\sigma^2}{\mu^2} + 1 = \frac{\Gamma\left(1+\frac{2}{\alpha}\right)}{\Gamma\left(1+\frac{1}{\alpha}\right)^2}$$

can be used to reduce α and β.

Weibull Distribution for Various Choices of α and β.





Description / Name

Probability Density Function

Comment

Lognormal Prob. Distribution

If the random variable **x** has a normal distribution with mean μ and standard deviation σ, the random variable **y** = e^x is said to have a lognormal distribution with parameters μ and σ. Therefore the random variable y of a lognormal distribution is continuous and ranges from 0 to ∞.

Mean:

$$E(y) = e^{\mu + \sigma^2/2}$$

Variance:

$$V(y) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$$

Median: e^{μ}

Mode: $e^{\mu - \sigma^2}$

CDF: $\Phi\left(\frac{\ln(x) - \mu}{\sigma}\right)$

where Φ is the CDF of the standard normal distribution

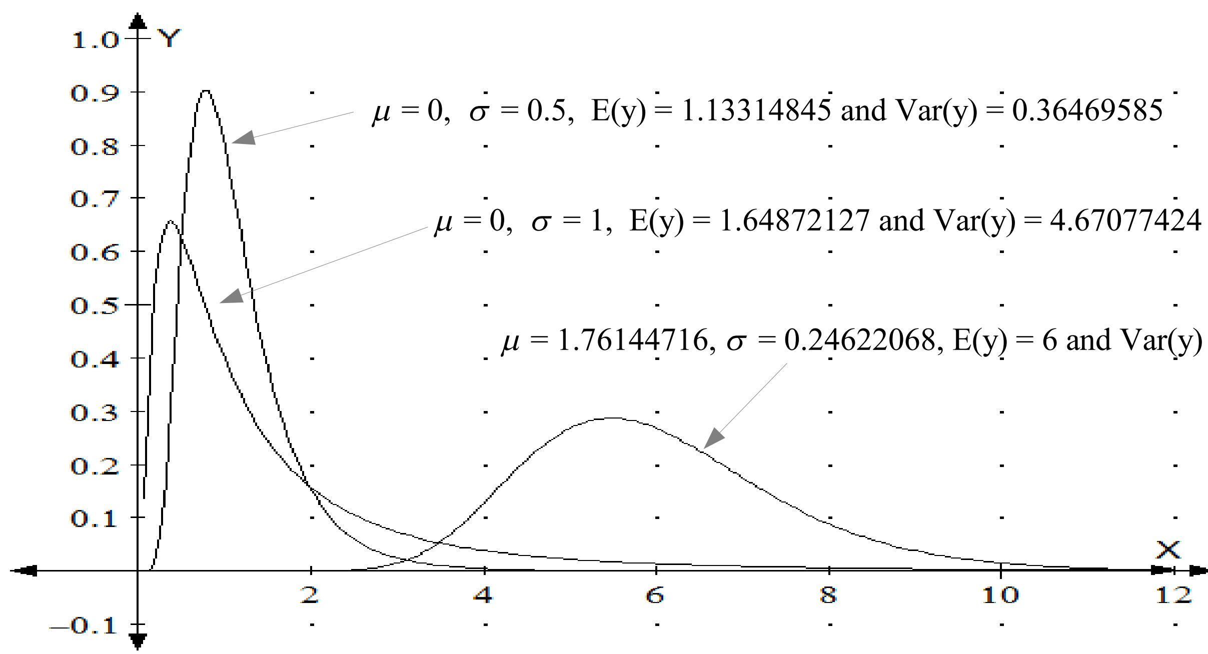
$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\left(\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right)} \quad \text{if } x > 0$$

$$= 0 \quad \text{if } x \le 0$$

If E(y) and V(y) are known, μ and σ^2 can be computed as follows:

$$\sigma^2 = \ln\left(1 + \frac{V(y)}{E(y)^2}\right)$$

$$\mu = \ln(E(y)) - \frac{1}{2}\sigma^2$$



The lognormal distribution is highly skewed to the right. This is the reason the lognormal distrubution is used to model processes that tend to produce occasional large values, or outliers.

Description / Name | **Probability Density Function** | **Expected Value** | **Standard Deviation** | **Comment**

Gamma Probability Distribution

The random variable **x** of a Gammal distribution is continuous and ranges from 0 to ∞ .

The properties of the distribution are completely determined by two positive real number constants, r and λ.

If r = 1, the gamma distribution is exponential with parameter λ.

If r is an integer, the gamma distribution is an extension of the exponential distribution where the random variable x equals the sum of r independent exponential random variables.

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} \quad \text{if } x > 0$$

$$= 0 \quad \text{if } x \le 0$$

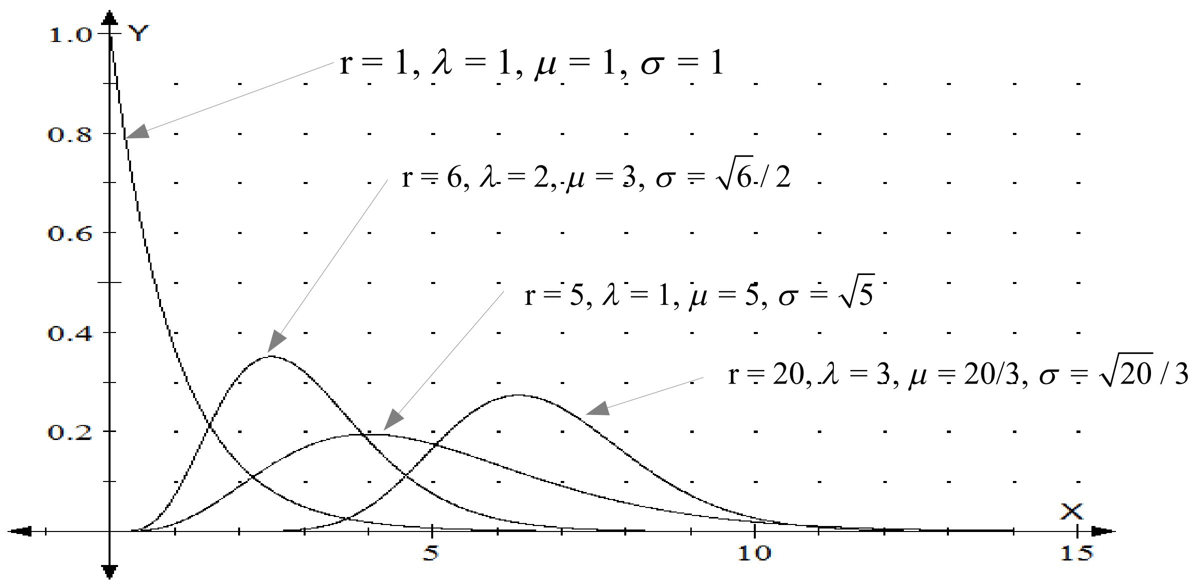
$$\mu = \frac{r}{\lambda} \quad \text{and} \quad \sigma = \frac{\sqrt{r}}{\lambda}$$

If r is an integer, then

$$F(x) = 1 - \sum_{k=1}^{r-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!}$$

One use of the gamma distribution is to model the waiting time until the r th success after some starting time in a Poisson process. The parameter λ of the Poisson distribution equals the mean number of successes or events in an interval of time.

Example: Suppose arrivals at a drive-through window follow a Poisson process with a mean arrival rate λ = 0.2 arrivals per minute. If we are interested in the probability distribution of waiting times until the 3rd arrival after some start time, we can use a gamma distribution with parameters r = 3 and λ = 0.2 to model the problem. 1/λ = 5 minutes per one arrival which equals the mean waiting time between events or successes in a Poisson process. Therefore μ = r/λ = 15 minutes which indicates that the mean arrival time until the 3 rd event in a Poisson process.



Description / Name	Probability Density Function	Expected Value	Standard Deviation
Negative Binomial Distribution	$P(x) = {}_{x+r-1}C_x \cdot (1-p)^r \cdot p^x$	$\mu = \dfrac{pr}{1-p}$	$\sigma = \dfrac{\sqrt{pr}}{1-p}$

A Bernoulli trial is an experiment that has two possible outcomes, named success or failure. The term 'success' does **not** necessarily mean that one of the outcomes is better than the other outcome; it is only a label that refers to one of the two possible outcomes.

p = the probability of success on each trial
q = the probability of failure on each trial.
r = a fixed constant which represents a number of failures. **r** > 0
x = the number of successes before **r** failures occurs.
x = 0, 1, 2, 3, 4, . . .

P(x) = the probability of having exactly **x** successes before **r** failure occurs.