Session 2
$\mathbb{Q}$uantitative Analysis of Financial Markets
Distributions

Christopher Ting

http://www.mysmu.edu/faculty/christophert/

Christopher Ting

✉: christophert@smu.edu.sg

☎: 6828 0364

LKCSB 5036

October 12, 2018

Broad Lesson Plan

Introduction

- It is important to distinguish between discrete distribution versus continuous distribution.
- Probability distributions can be divided into two broad categories:
  - parametric distributions, which is described by a mathematical function
  - nonparametric distributions, which are not described by a mathematical formula
- A major advantage of dealing with nonparametric distributions is that assumptions required are minimum. "Data speak for themselves."

Learning Outcomes of QA03

Chapter 4.

Michael Miller, Mathematics and Statistics for Financial Risk Management, 2nd Edition.

(Hoboken, NJ: John Wiley & Sons, 2013).

- Distinguish the key properties and identify common occurrences of each distributions.
- Discrete
  - Bernoulli distribution
  - Binomial distribution
  - Poisson distribution
- Continuous Piecewise Linear
  - Uniform distribution
  - Triangular distribution
- Continuous and differential
  - Normal distribution
  - Student's t distribution
  - Lognormal distribution
  - Chi-squared distribution
  - F distribution

Bernoulli Distribution

- A Bernoulli random variable X is equal to either zero or one.
- If we define p as the probability that X equals one, we have

$$\mathbb{P}\left(X = 1\right) = p; \qquad \mathbb{P}\left(X = 0\right) = 1 - p.$$

- Binary outcomes are common in finance: a bond can default or not default; the return of a stock can be positive or negative; a central bank can decide to raise rates or not to raise rates.
- The mean and variance of a Bernoulli random variable are

$$\mu = p \times 1 + (1 - p) \times 0 = p$$
$$\sigma^2 = p \times (1 - p)^2 + (1 - p) \times (0 - p)^2 = p(1 - p).$$

- Bernoulli Distribution is discrete, and parametric with p being the only parameter.

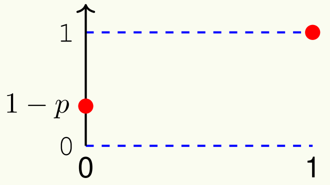
More on Bernoulli Distribution

- The probability of a Bernoulli random variable can be written as

$$\mathbb{P}(X = x) = p^x(1-p)^{1-x}.$$

- The cumulative distribution function is

$$F(X = x) = \begin{cases} 1-p, & \text{for } x = 0; \\ 1, & \text{for } x = 1. \end{cases}$$



Frequentist Estimation

- How can we estimate p of a coin for the head to turn up?
- Throw N times, and count the number (n) of times the head has turned up. Then
$$\widehat{p} = \frac{n}{N}.$$
- Is this estimator unbiased?
- To answer this question, we need the binomial distribution.

Binomial Distribution

- The binomial distribution gives the discrete probability distribution $\mathbb{P}(n, N; p)$ of obtaining exactly n successes out of N Bernoulli trials.
- The binomial distribution is given by

$$\mathbb{P}(n; N, p) = \binom{N}{n} p^n (1-p)^{N-n}.$$

- The cumulative distribution function is

$$F(k; N, p) = \mathbb{P}(X \leq k) = \sum_{i=0}^{\lfloor k \rfloor} \binom{N}{i} p^i (1-p)^{N-i},$$

where the symbol $\lfloor k \rfloor$ is the largest integer smaller than k.

Unbiasedness of $\widehat{p}$

- The probability of obtaining n successes in N trials is the binomial probability: $\mathbb{P}(n; N, p)$.
- The expected value of the estimator $\widehat{p} = n/N$ is therefore given by

$$
\begin{aligned}
\mathbb{E}\left(\widehat{p}\right) &= \sum_{n=0}^{N} \frac{n}{N} \binom{N}{n} p^n (1-p)^{N-n} = p \sum_{n=1}^{N} \frac{N!}{(N-n)!n!} \frac{n}{N} p^{n-1} (1-p)^{N-n} \\
&= p \sum_{n=1}^{N} \frac{(N-1)!}{(N-n)!(n-1)!} p^{n-1} (1-p)^{(N-1)-(n-1)} \\
&= p(1-p)^{N-1} \sum_{m=0}^{N-1} \binom{N-1}{m} \left(\frac{p}{1-p}\right)^m \\
&= p(1-p)^{N-1} \left(1 + \frac{p}{1-p}\right)^{N-1} = p(1-p)^{N-1} \left(\frac{1}{1-p}\right)^{N-1} = p.
\end{aligned}
$$

Mean of Binomial Distribution

- Consider the moment generating function $M(t)$, which is the expected value of e^{tn} with t being the (auxiliary) variable.

$$
\begin{aligned}
M(t) = \mathbb{E}\left(e^{tn}\right) &= \sum_{n=0}^{N} e^{tn} \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=0}^{N} \binom{N}{n} \left(pe^t\right)^n (1-p)^{N-n} \\
&= \left(pe^t + (1-p)\right)^N
\end{aligned}
$$

- The first derivative of $M(t)$ with respect to t is

$$
M'(t) = N\left(pe^t + (1-p)\right)^{N-1}\left(pe^t\right).
$$

The mean is $\mu = M'(0) = N(p + 1 - p)p = Np$.

Variance of Binomial Distribution

- The second derivative is

$$M''(t) = N(N-1)\big(pe^t+(1-p)\big)^{N-2}\big(pe^t\big)^2+N\big(pe^t+(1-p)\big)^{N-1}\big(pe^t\big).$$

- At $t = 0$, $\mathbb{E}\left(X^2\right) = M''(0) = N(N-1)p^2 + Np$.
- Since $\mathbb{V}\left(X\right) = \mathbb{E}\left(X^2\right) - \mu^2$, we find that

$$\mathbb{V}\left(X\right) = Np(1-p).$$

- What is your intuition for mean being Np and variance being $Np(1-p)$?

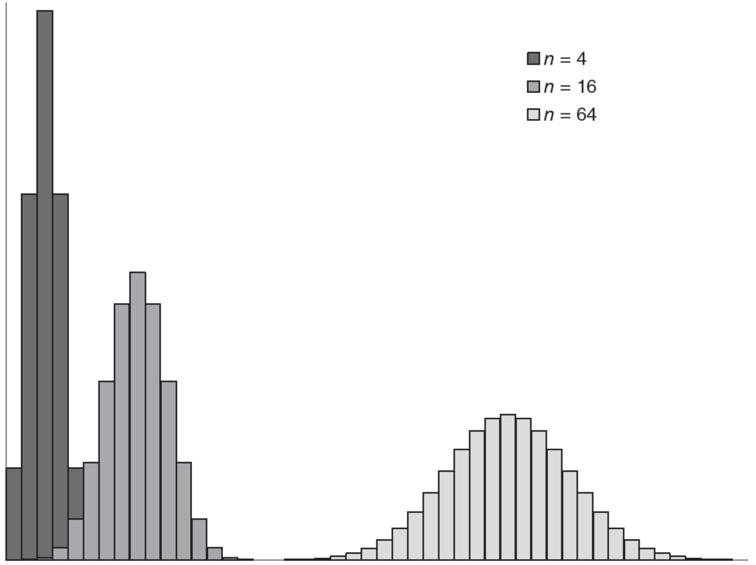
Sample Problem

- Suppose you have four bonds, each with a 10% probability of defaulting over the next year. The event of default for any given bond is independent of the other bonds defaulting. What is the probability that zero, one, two, three, or all of the bonds default? What is the mean number of defaults? The standard deviation?

- Answer: We calculate the probability of each possible outcome as follows:

Defaults	$\binom{n}{k}$	$p^k(1-p)^{n-k}$	Probability
0	1	65.61%	65.61%
1	4	7.29%	29.16%
2	6	0.81%	4.86%
3	4	0.09%	0.36%
4	1	0.01%	0.01%

Binomial Probability Mass Function



Poisson Distribution

- For a Poisson random variable X,

$$\mathbb{P}\left(X = n\right) = \frac{\lambda^n}{n!} e^{-\lambda},$$

  where λ is the constant parameter.

- The Poisson distribution is often used to model
  1. the occurrence of events over time such as the number of bond defaults in a portfolio
  2. the number of crashes in equity markets.
  3. jumps in jump-diffusion models

Poisson Distribution from Binomial Model

- Recall that the probability of obtaining exactly n successes in N Bernoulli trials is given by

$$\mathbb{P}\left(X = n; N, p\right) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}.$$

- Viewing the distribution as a function of the expected number of successes $\lambda := Np$ instead of the sample size N for fixed p, we write

$$\mathbb{P}\left(X = n; N, \lambda := Np\right) = \frac{N!}{n!(N-n)!} \left(\frac{\lambda}{N}\right)^n \left(1 - \frac{\lambda}{N}\right)^{N-n}.$$

Poisson Distribution from Binomial Model (cont'd)

- What happens if the sample size N becomes large, even infinite?

$$
\begin{aligned}
P_\lambda(n) &:= \lim_{N\to\infty} \mathbb{P}\left(X = n; N, \lambda := Np,\right) \\
&= \lim_{N\to\infty} \frac{N(N-1)\cdots(N-n+1)}{n!} \frac{\lambda^n}{N^n} \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^{-n} \\
&= \lim_{N\to\infty} \frac{N(N-1)\cdots(N-n+1)}{N^n} \frac{\lambda^n}{n!} \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^{-n} \\
&= 1 \times \frac{\lambda^n}{n!} e^{-\lambda} \times 1 \\
&= \frac{\lambda^n}{n!} e^{-\lambda}
\end{aligned}
$$

- Poisson distribution is a limiting case of binomial model.

Mean of Poisson Random Variable

- First we show that $P_\lambda(n)$, $n = 0, 1, 2, \ldots, \infty$ indeed adds up to 1.

$$\sum_{n=0}^{\infty} P_\lambda(n) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1.$$

- The mean is

$$\sum_{n=0}^{\infty} n P_\lambda(n) = e^{-\lambda} \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} = e^{-\lambda} \lambda \sum_{n=0}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} = \lambda.$$

Variance of Poisson Random Variable

- To compute variance, we first compute, knowing that $n^2 = n(n-1) + n$,

$$\sum_{n=0}^{\infty} n^2 P_\lambda(n) = e^{-\lambda}\lambda^2 \sum_{n=0}^{\infty} \frac{\lambda^{n-2}}{(n-2)!} + e^{-\lambda}\lambda \sum_{n=0}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} = \lambda^2 + \lambda.$$

- Hence the variance is $\lambda^2 + \lambda - \lambda^2 = \lambda$.

Poisson Process

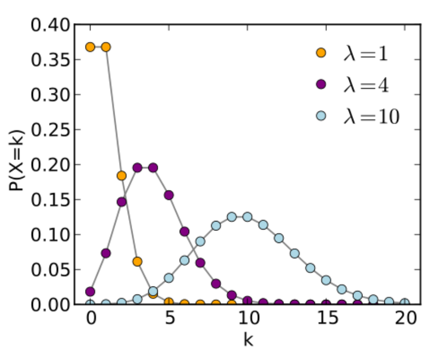
- If the rate at which events occur over time is constant, and the probability of any one event occurring is independent of all other events, then we say that the events follow a Poisson process:

$$\mathbb{P}\big(X := N(t+\tau) - N(t) = n\big) = \frac{(\lambda\tau)^n}{n!}e^{-\lambda\tau}.$$

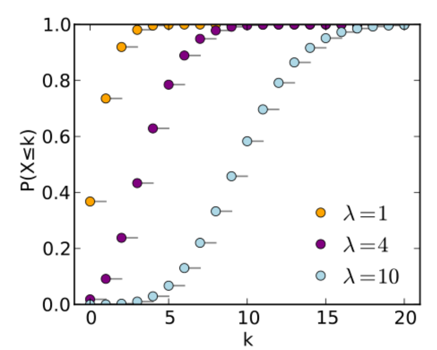
  Here, $N(t+\tau) - N(t) = n$ is the number of events in the time interval $(t, t+\tau]$.

- The mean and variance of a Poisson process are the same: λt.

Probability Mass Function and CDF



Probability Mass Function



Cumulative Distribution Function

Source: https://en.wikipedia.org/

Sample Problem

- Assume that defaults in a large bond portfolio follow a Poisson process. The expected number of defaults each month is four. What is the probability that there are exactly three defaults over the course of one month? Over two months?
- Over one month, the probability is

$$\mathbb{P}\left(X = 3\right) = \frac{(\lambda\tau)^n}{n!}e^{-\lambda\tau} = \frac{(4\times 1)^3}{3!}e^{-4\times 1} = 19.5\%.$$

- Over two months, the probability is

$$\mathbb{P}\left(X = 3\right) = \frac{(\lambda\tau)^n}{n!}e^{-\lambda\tau} = \frac{(4\times 2)^3}{3!}e^{-4\times 2} = 2.9\%.$$

Uniform Distribution

- So far, we have been talking about discrete distributions.
- The continuous uniform distribution is one of the most fundamental distributions in statistics.
- For all $x \in [b_1, b_2]$,
$$u(b_1, b_2) = c.$$
In other words, the probability density is constant and equal to c between b_1 and b_2, and zero everywhere else.
- Because the probability of any outcome occurring must be one, you can find the value of c as ____________________.

Mean, Variance, PDF, and CDF

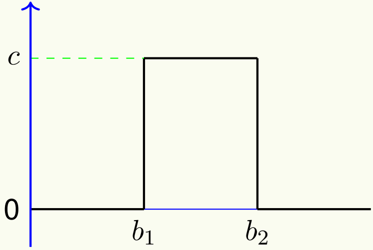
- The mean of the uniform random variable is

  ______________________.

- The variance of the uniform random variable is

  ______________________.

- Probability Density Function

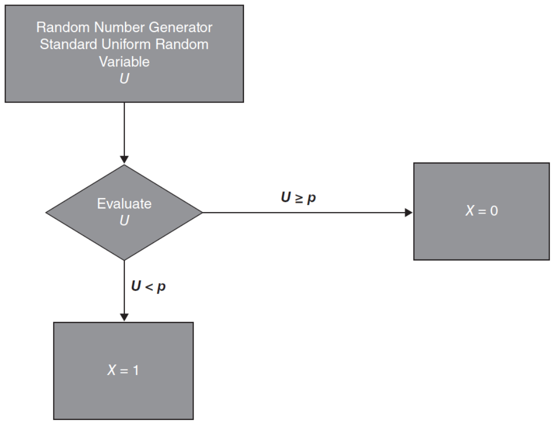


- The cumulative distribution function is

  ______________________.

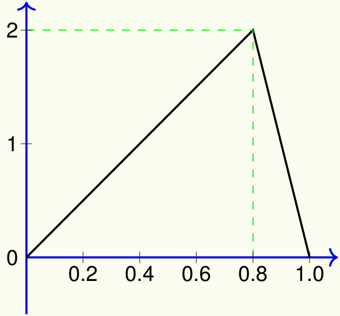
Application of Uniform Random Variable

In a computer simulation, one way to model a Bernoulli variable is to start with a standard uniform variable, which is when $b_1 = 0$ and $b_2 = 1$.



Triangular Distribution

Example



- The PDF for a triangular distribution with a minimum of a, a maximum of b, and a mode of c is described by the following two-part function:

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)}, & x \in [a, c] \\ \dfrac{2(b-x)}{(b-a)(b-c)}, & x \in (c, b] \end{cases}$$

- When modeling default rates and recovery rates, triangular distributions are useful.
- **Class Exercise**: What are the values of a, b, and c for the example?

Normal Distribution

- Normal distribution is also referred to as the **Gauss**ian distribution.
- For a random variable X, the probability density function for the normal distribution is

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

and we write

$$X \sim N(\mu,\sigma^2),$$

which means "X is normally distributed with mean μ and variance σ^2."

Properties of Normal Random Variable

- Any linear combination of *independent* normal random variables is also normal.

$$Z = aX + bY \sim N\left(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2\right).$$

- Example: If the log returns of individual stocks are independent and normally distributed, then the average return of those stocks will also be normally distributed.
- The bell shape curve has 0 skewness and kurtosis of 3.

Standard Normal Random Variable

- Standard normal distribution is $N(0,1)$ with the probability density function:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

- Because a linear combination of independent normal distributions is also normal, standard normal distributions are the building blocks of many financial models.

- To get a normal variable with a standard deviation of σ and a mean of μ, we simply multiply the standard normal variable φ by σ and add μ.

$$X = \mu + \sigma\varphi \sim N(\mu, \sigma^2).$$

Two Correlated Normal Variables

To create two correlated normal variables, we can combine three independent standard normal variables, X_1, X_2, and X_3, as follows:

$$X_A = \sqrt{\rho} X_1 + \sqrt{1-\rho} X_2$$

$$X_B = \sqrt{\rho} X_1 + \sqrt{1-\rho} X_3$$

Class Exercise: Show that X_A and X_B are standard normal random variables.

Class Exercise: Show that X_A and X_B are indeed correlated with correlation coefficient ρ.

Log Return and Normal Distribution

- Normal distributions are used throughout finance and risk management.
- Normally distributed log returns are widely used in financial simulations as well, and form the basis of a number of financial models, including the Black-Scholes option pricing model.
- One attribute that makes log returns particularly attractive is that they can be modeled using normal distributions.
  1. A normal random variable can realize values ranging from $-\infty$ to ∞.
  2. Simple return has a minimum, -100%.
  3. But log return does not have a minimum as it can potentially be $-\infty$, and thus more amenable to modeling by a normal distribution.

Normal Distribution Confidence Intervals

	One-Tailed	Two-Tailed
1.00%	-2.33	-2.58
2.50%	-1.96	-2.24
5.00%	-1.64	-1.96
10.00%	-1.28	-1.64
90.00%	1.28	1.64
95.00%	1.64	1.96
97.50%	1.96	2.24
99.00%	2.33	2.58

- The normal distribution is symmetrical, 5% of the values are less than 1.64 standard deviations below the mean.
- The two-tailed value tells you that 95% of the mass is within ± 1.96 standard deviations of the mean. So, ____% of the outcomes are less than -1.96 standard deviations from the mean.

Lognormal Distribution

- Can we use another distribution and model the standard simple returns directly?
- Yes, use the lognormal distribution instead. For $x > 0$,

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2}.$$

- If a variable has a lognormal distribution, then the log of that variable has a normal distribution. So, if log returns are assumed to be normally distributed, then one plus the simple return will be lognormally distributed.
- Using the lognormal distribution provides an easy way to ensure that returns less than -100% are avoided.

More on Lognormal Distribution

- The lognormal pdf can also be written as

$$f(x) = e^{\frac{1}{2}\sigma^2-\mu}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln x-(\mu-\sigma^2)}{\sigma}\right)^2}.$$

  It shows that the lognormal distribution is asymmetrical and peaks at $\exp\left(\mu-\sigma^2\right)$.

- Given μ and σ, the mean and variance are given by

$$\mathbb{E}\left(X\right) = e^{\mu+\frac{1}{2}\sigma^2}$$
$$\mathbb{V}\left(X\right) = \left(e^{\sigma^2}-1\right)e^{2\mu+\sigma^2}.$$

- When it comes to modeling, it is often easier to work with log returns and normal distributions than with standard returns and lognormal distributions.

Chi-Squared Distribution

- If we have k independent standard normal variables, $Z_1, Z_2, \ldots, Z_k$, then the sum of their squares, S, has a chi-squared distribution.

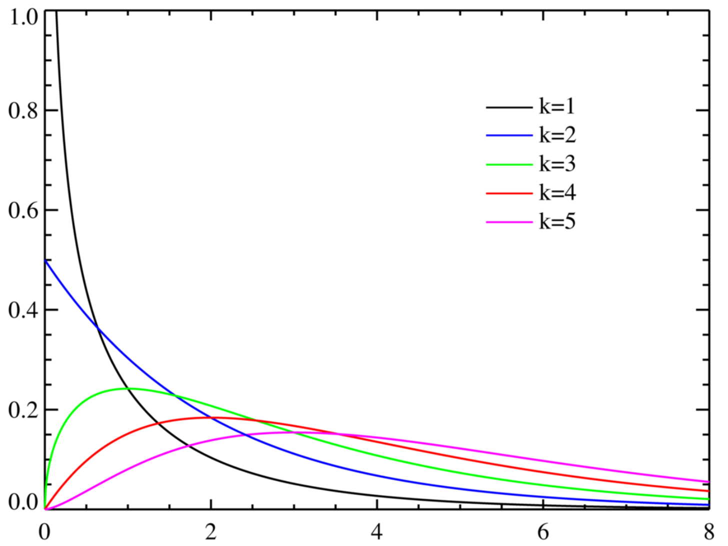
$$S := \sum_{i=1}^{k} Z_i^2; \qquad S \sim \chi_k^2$$

- The variable k is commonly referred to as the degrees of freedom.
- For positive values of x, the probability density function for the chi-squared distribution is

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},$$

where Γ is the gamma function: $\Gamma(n) := \int_0^\infty x^{n-1} e^{-x}\, dx.$

Chi-Square Probability Density Function



More on Chi-Squared Distribution

- The mean of the distribution is k, and the variance is $2k$.

- As k increases, the chi-squared distribution becomes increasingly symmetrical. As k approaches infinity, the chi-squared distribution converges to the normal distribution.

- The chi-squared distribution is widely used in risk management, and in statistics in general, for hypothesis testing.

Student's t Distribution

- **William Sealy Gosset**, an employee at the Guinness brewery in Dublin, used the pseudonym Student for journal publications.
- If Z is a standard normal variable and U is a chi-square variable with k degrees of freedom, which is independent of Z, then the random variable X,

$$X = \frac{Z}{\sqrt{U/k}},$$

  follows a t distribution with k degrees of freedom.
- The probability density function of Student's t random variable is

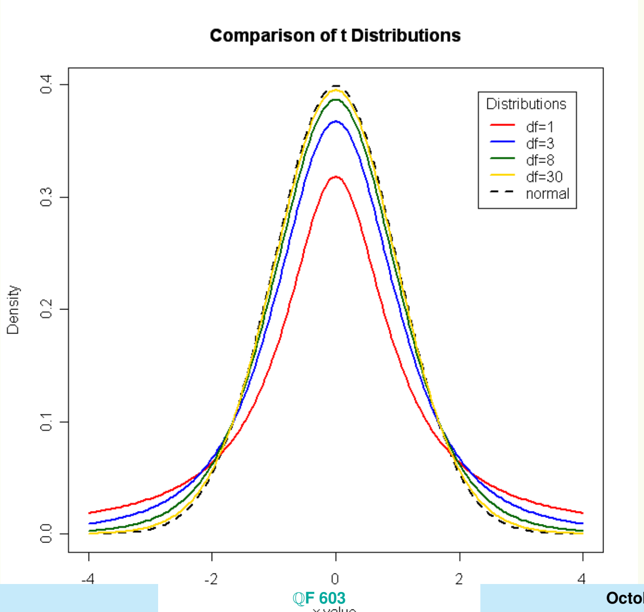
$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}},$$

  where k is the degrees of freedom and $\Gamma(x)$ is the gamma function.

Characteristics of Student's t Distribution

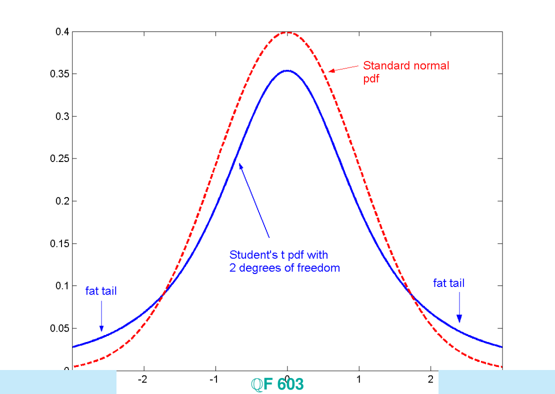
- The t distribution is symmetrical around its mean, which is equal to zero.
- For low values of k, the t distribution looks very similar to a standard normal distribution, except that it displays excess kurtosis.
- As k increases, this excess kurtosis decreases. In fact, as k approaches infinity, the t distribution converges to a standard normal distribution.
- The variance of the t distribution for $k > 2$ is $\dfrac{k}{k-2}$. As k increases, the variance of the t distribution converges to one, the variance of the standard normal distribution.

Comparison of t Distributions



Probability Density Function of t Statistics

- The pdf of t statistics have “fatter” tails compared to standard normal.
- When the number of degrees of freedom is ∞, Student’s t pdf becomes standard normal pdf.



Chi-Square and Student t Distributions

- If random variable $Z \stackrel{d}{\sim} N(0,1)$, then

$$V := Z^2 \stackrel{d}{\sim} \chi^2_1.$$

- More generally,

$$\sum_{i=1}^{n} Z_i^2 \stackrel{d}{\sim} \chi^2_n.$$

- If $X \stackrel{d}{\sim} N(0,1)$, and $V \stackrel{d}{\sim} \chi^2_n$, and both X, V are independent, then the following statistic follows Student's t distribution with n degrees of freedom, i.e.,

$$\frac{X}{\sqrt{\frac{V}{n}}} \stackrel{d}{\sim} t_n.$$

F-Distribution

- If U_1 and U_2 are two independent chi-squared distributions with k_1 and k_2 degrees of freedom, respectively, then X,

$$X = \frac{U_1/k_1}{U_2/k_2} \sim F(k_1, k_2)$$

follows an F-distribution with parameters k_1 and k_2.

- The probability density function of the F-distribution is

$$f(x) = \frac{\sqrt{\frac{(k_1 x)^{k_1} k_2^{k_2}}{\left(k_1 x + k_2\right)^{k_1+k_2}}}}{xB\left(k_1/2, k_2/2\right)},$$

where $B(x, y)$ is the beta function:

$$B(x, y) = \int_0^1 z^{x-1}(1-z)^{y-1}\, dz.$$

A Bit More on F-Distribution

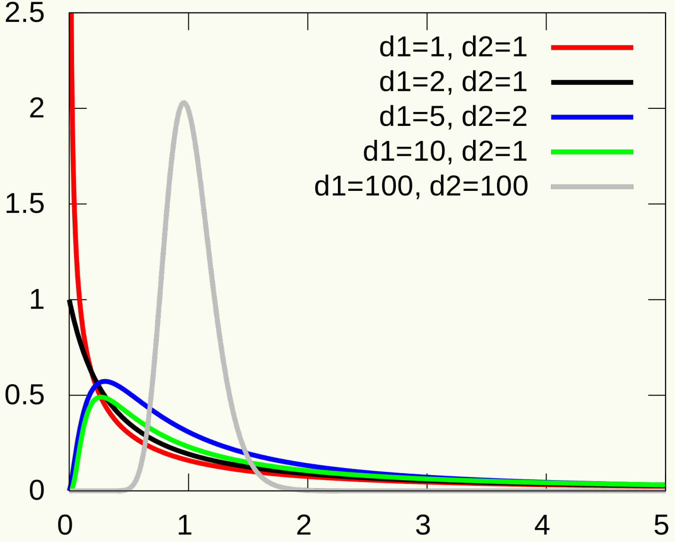
- The mean and variance of the F-distribution are as follows

$$\mu = \frac{k_2}{k_2 - 2}, \qquad \text{for } k_2 > 2,$$
$$\sigma^2 = \frac{2k_2^2\big(k_1 + k_2 - 2\big)}{k_1\big(k_2 - 2\big)^2(k_2 - 4)}, \qquad \text{for } k_2 > 4.$$

- As k_1 and k_2 increase to infinity, the mean and mode converge to one, and the F-distribution converges to a normal distribution.
- The square of a variable with a t distribution has an F-distribution. More specifically, if X is a random variable with a t distribution with k degrees of freedom, then X^2 has an F-distribution with 1 and k degrees of freedom:

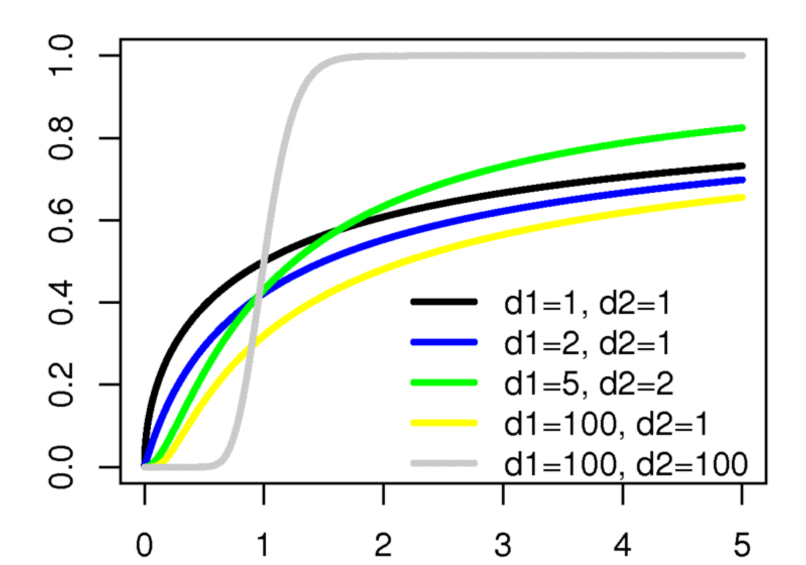
$$X^2 \sim F(1, k).$$

F Probability Density Function



CDF of F Distributions

(source: http://en.wikipedia.org/wiki/File:F_distributionCDF.png)



Sample Variance and Chi-Square Statistic

- Let $Z_1, Z_2, \ldots, Z_k$ be k independent random variables, and each Z is a standard normal random variable. The sum of the squared Zs also follows a chi-squared distribution with k degrees of freedom.

$$\sum_{i=1}^{k} Z_i^2 \overset{d}{\sim} \chi_k^2.$$

- The unbiased sample variance s^2 is a random variable owing to random sampling.

$$\sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sigma} \right)^2 = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2}{\sigma^2} = (n-1)\frac{s^2}{\sigma^2} \overset{d}{\sim} \chi_{n-1}^2.$$

Binomial Random Variable

- There are k bins. The probability that a data point falls in bin i is p_i. If the data are drawn independently from each other, then the count (number of occurrences) O_i for bin i is a binomial random variable with probability distribution

$$\mathbb{P}(O_i = j) = \binom{n}{j} p_i^j (1 - p_i)^{n-j}.$$

- Mean and variance of O_i are, respectively,

$$\mu_i = np_i \qquad \text{and} \qquad \sigma_i^2 = np_i - np_i^2$$

- Note that $p_i < 1$, so $\sigma_i^2 \approx np_i = \mu_i$. So the standard normal random variable Z_i can be approximated by

$$Z_i = \frac{X_i - \mu_i}{\sigma} \approx \frac{X_i - \mu_i}{\sqrt{\mu_i}} \overset{d}{\sim} N(0, 1).$$

Chi-Square Test

- Moreover,

$$\sum_{i=1}^{k} Z_i^2 = \sum_{i=1}^{k} \frac{\left(X_i - \mu_i\right)^2}{\mu_i} \overset{d}{\sim} \chi^2_{k-1}.$$

- Example: There are 1500 returns and 55 returns are smaller than the fifth percentile. What is the chi-square statistic?
- Solution:
  1. There are only two bins: either below or above the fifth percentile.
  2. Fifth percentile means that there is a 5% chance of falling into the "below" bin.
  3. So $\mu_1 = 1500 \times 0.05 = 75$, and $\mu_2 = 1500 \times 0.95 = 1425$

$$\frac{(55-75)^2}{75} + \frac{\left((1500-55)-1425\right)^2}{1425} = 5.61.$$

Chi-Square Test: General Case

Issues	NYSE	Nasdaq	NYSE MKT	Total
Advances	2,164	1,794	228	4,186
Declines	987	997	150	2,134
Unchanged	102	69	27	198
Total	3,253	2,860	405	6,518

Source: Wall Street Journal

1. Set up hypotheses and determine level of significance.

H_0 : stock movement and exchange listing are independent.

H_1 : H_0 is false. $\alpha = 0.05$

2. Compute the expected frequency

$$\text{Expected Frequency}_{ij} = \frac{\text{Row Total}_i \times \text{Column Total}_j}{\text{Grand Total}}.$$

Real-Life Example: Results

Issues	NYSE	Nasdaq	NYSE MKT	Total
Advances	2,089.1	1,836.8	260.1	4,186.0
Declines	1,065.0	936.4	132.6	2,134.0
Unchanged	98.8	86.9	12.3	198.0
Total	3,253.0	2,860.0	405.0	6,518.0

Expected Frequencies

$$\begin{aligned}\chi^2 =& \frac{(2164-2089.1)^2}{2089.1}+\frac{(1794-1836.8)^2}{1836.8}+\frac{(228-260.1)^2}{260.1}\\ &+\frac{(987-1065.0)^2}{1065.0}+\frac{(997-936.4)^2}{936.4}+\frac{(150-132.6)^2}{132.6}\\ &+\frac{(102-98.8)^2}{98.8}+\frac{(69-86.9)^2}{86.9}+\frac{(27-12.3)^2}{12.3}=40.91.\end{aligned}$$

Degrees of Freedom = (rows -1) $\times$ (columns-1) = 4.

Summary

F_{n_1,n_2} — $F_{n_1,n_2} = \dfrac{\chi^2_{n_1}/n_1}{\chi^2_{n_2}/n_2}$

Student's t_n — $\lim_{n\to\infty} t_n \longrightarrow Z$

Chi square χ^2_n — $V := \sum_{i=1}^{n} Z_i^2 \overset{d}{\sim} \chi^2_n$

Standard normal Z — $Z = \dfrac{r-\mu}{\sigma}$

Takeaways

- All the 10 distributions are parametric, being dependent on parameters that can be interpreted intuitively.
- Additionally, Student's t, χ^2, and F distributions need degrees of freedom to determine their shape.
- Probability mass function is the discrete analogue of probability density function.
- χ^2 is the sum of k squared of independent standard normal variables.
- t variable is made of a standard normal random variable and a χ^2 random variable, which are independent.
- The F random variable is the ratio of two independent χ^2 random variables.