

Unit 5: Sampling Distributions of Statistics

Statistics 571: Statistical Methods
Ramón V. León

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

1

Definitions and Key Concepts

- A sample statistic used to estimate an unknown population parameter is called an **estimate**
- The discrepancy between the estimate and the true parameter value is known as **sampling error**
- Sampling error is due to sampling variation

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

2

Frequentist Approach to Statistics

- Assesses the accuracy of a sample estimate by considering how the estimate would vary around the true parameter value if repeated random samples are drawn from the same population
- A statistic is a random variable with a probability distribution - called the **sampling distribution** - which is generated by repeated sampling.
- We use the sampling distribution of a statistic to assess the sampling error of an estimate

Sample Mean

A **random sample** is a set of independently, identically distributed or i.i.d. observations X_1, X_2, \dots, X_n (when sampling from a large population or with replacement)

Assume that the population has

mean $\mu = E(X_i)$ and variance $\sigma^2 = Var(X_i)$

How does the sample mean $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

vary on repeated random samples of size n ?

This is called the sampling distribution of the sample mean.

Mean and Variance of a Die Toss

$$\mu = \frac{1}{6}(1 + 2 + \dots + 6) = 3.5$$

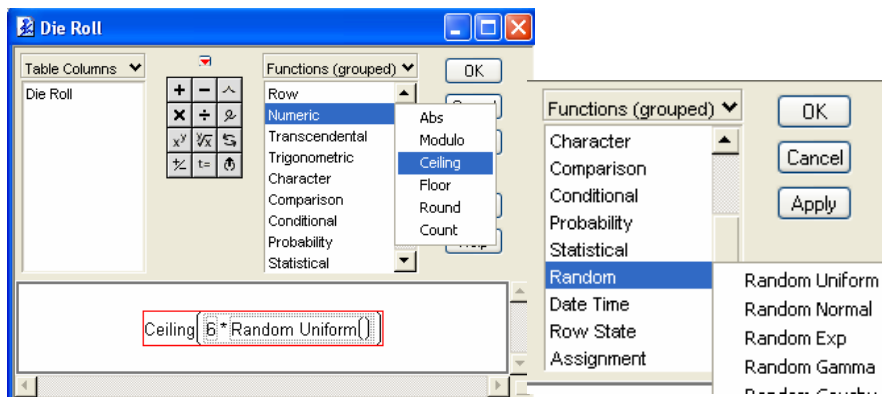
$$\sigma^2 = E(X^2) - \mu^2 = \frac{1}{6}(1^2 + 2^2 + \dots + 6^2) - (3.5)^2 = 2.9166.$$

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

5

Simulating a Die Toss in JMP



6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

6

Rolling Two Dice

Each one of these 36 outcomes are equally likely, i.e., each one occurs with 1/36 probability.

Table 5.1 All Possible Samples of Size 2 from a Uniform Distribution on 1, 2, . . . , 6 and Associated Sample Means and Variances

(x_1, x_2)	\bar{x}	s^2	(x_1, x_2)	\bar{x}	s^2
(1, 1)	1.0	0.0	(4, 1)	2.5	4.5
(1, 2)	1.5	0.5	(4, 2)	3.0	2.0
(1, 3)	2.0	2.0	(4, 3)	3.5	0.5
(1, 4)	2.5	4.5	(4, 4)	4.0	0.0
(1, 5)	3.0	8.0	(4, 5)	4.5	0.5
(1, 6)	3.5	12.5	(4, 6)	5.0	2.0
(2, 1)	1.5	0.5	(5, 1)	3.0	8.0
(2, 2)	2.0	0.0	(5, 2)	3.5	4.5
(2, 3)	2.5	0.5	(5, 3)	4.0	2.0
(2, 4)	3.0	2.0	(5, 4)	4.5	0.5
(2, 5)	3.5	4.5	(5, 5)	5.0	0.0
(2, 6)	4.0	8.0	(5, 6)	5.5	0.5
(3, 1)	2.0	2.0	(6, 1)	3.5	12.5
(3, 2)	2.5	0.5	(6, 2)	4.0	8.0
(3, 3)	3.0	0.0	(6, 3)	4.5	4.5
(3, 4)	3.5	0.5	(6, 4)	5.0	2.0
(3, 5)	4.0	2.0	(6, 5)	5.5	0.5
(3, 6)	4.5	4.5	(6, 6)	6.0	0.0

Rolling Two Dice Sampling Distribution

Table 5.2 Sampling Distribution of \bar{X} for Samples of Size 2 from a Uniform Distribution on 1, 2, . . . , 6

\bar{x}	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0
$f(\bar{x})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$E(\bar{X}) = 1.0 \times \frac{1}{36} + 2.0 \times \frac{2}{36} + \dots + 6.0 \times \frac{1}{36} = 3.5$$

and

$$\begin{aligned} \text{Var}(\bar{X}) &= (1.0 - 3.5)^2 \times \frac{1}{36} + (2.0 - 3.5)^2 \times \frac{2}{36} + \dots + (6.0 - 3.5)^2 \times \frac{1}{36} \\ &= 1.4583 = 2.9166/2. \end{aligned}$$

These values are consistent with the formulas

$$E(\bar{X}) = \mu \quad \text{and} \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

Homework To be Done Right Away

Use the **Sampling Distribution simulation** Java applet at the Rice Virtual Lab in Statistics to do the following.

- Draw 10,000 random samples of size $N=5$ from the normal distribution provided.
 - Construct the histogram of the sampling distribution of the sample mean.
 - Construct the histogram of the sampling distribution of the sample variance
 - Turn in this output with the rest of the homework for Unit 5.
- Draw 10,000 random samples of size $N=20$ from the normal distribution provided.
 - Construct the histogram of the sampling distribution of the sample mean.
 - Construct the histogram of the sampling distribution of the sample variance
- Draw 10,000 random samples of size $N=5$ from a uniform distribution on $[0,32]$.
 - Construct the histogram of the sampling distribution of the sample mean.
 - Construct the histogram of the sampling distribution of the sample variance
- Draw 10,000 random samples of size $N=20$ from a uniform distribution on $[0,32]$.
 - Construct the histogram of the sampling distribution of the sample mean.
 - Construct the histogram of the sampling distribution of the sample variance
- Draw 10,000 random samples of size $N=5$ from the skewed distribution provided.
 - Construct the histogram of the sampling distribution of the sample mean.
 - Construct the histogram of the sampling distribution of the sample variance
 - Construct the histogram of the sampling distribution of the sample median
- Draw 10,000 random samples of size $N=20$ from the skewed distribution provided.
 - Construct the histogram of the sampling distribution of the sample mean.
 - Construct the histogram of the sampling distribution of the sample variance
 - Construct the histogram of the sampling distribution of the sample median

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

9

Distribution of Sample Means

- If the i.i.d. r.v.'s are Bernoulli, Normal, or Exponential the distribution of the sample mean can be calculated exactly.
- However, in general the exact distribution of the sample mean is difficult to calculate.
- What can be said about the distribution of the sample mean when the sample is drawn from an arbitrary population?
- In many cases we can approximate the distribution of the sample mean when n is large by a normal distribution. This result is called the **Central Limit Theorem**.

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

10

Central Limit Theorem

Let X_1, X_2, \dots, X_n be a random sample drawn from an arbitrary distribution with a finite mean μ and variance σ^2 . Then if n is sufficiently large

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$$

Sometimes the theorem is given in terms of the sums:

$$\frac{\sum_{j=1}^n X_j - n\mu}{\sigma\sqrt{n}} \approx N(0,1)$$

Central Limit Theorem Illustration

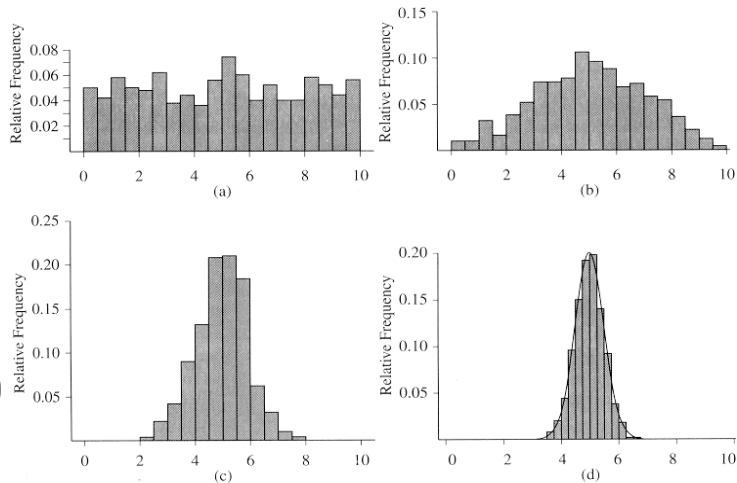
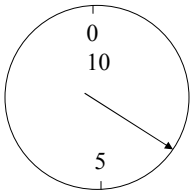
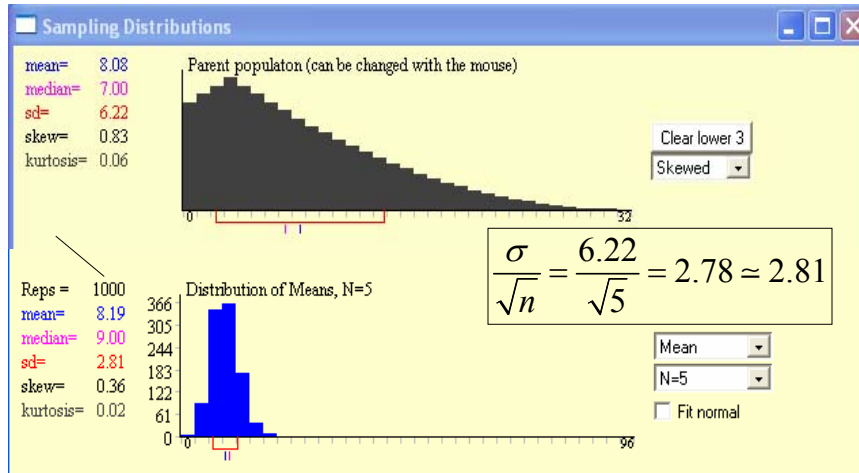


Figure 5.1 Histograms of Means of 500 Random Samples from a $U[0, 10]$ Distribution: (a) Single observations from a uniform distribution; (b) Sample means from $n = 2$ observations; (c) Sample means from $n = 10$ observations; (d) Sample means from $n = 30$ observations.

Screen Shots of the Output of the Sampling Distribution Simulation Java Applet



6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

13

Central Limit Theorem and Law of Large Numbers

- Both are asymptotic results about the sample mean
- Law of Large Numbers says that as n goes to infinity the sample mean converges to the population mean, i.e.

$$\bar{X} - \mu \text{ converges to } 0 \text{ as } n \rightarrow \infty$$

- CLT says that as n goes to infinity

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \text{ converges to } N(0,1) \text{ as } n \rightarrow \infty$$

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

14

Central Limit Theorem

Let X_1, X_2, \dots, X_n be a random sample drawn from an arbitrary distribution with a finite mean μ and variance σ^2 . Then if n is sufficiently large

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$$

Sometimes the theorem is given in terms of the sums:

$$\frac{\sum_{j=1}^n X_j - n\mu}{\sigma\sqrt{n}} \approx N(0,1)$$

Normal Approximation to the Binomial

A binomial r.v. is the sum of i.i.d. Bernoulli r.v.'s so the CLT can be used to approximate its distribution

Suppose that Z is Bernoulli. Then the mean of Z is p and its variance is $p(1-p)$.

By the CLT we have for the Binomial (n, p) r.v X :

$$\frac{X - np}{\sqrt{np(1-p)}} = \frac{\sum_{i=1}^n Z_i - np}{\sqrt{np(1-p)}} = \frac{\sum_{i=1}^n Z_i - nE(Z)}{\sqrt{Var(Z)} \times \sqrt{n}} \approx N(0,1)$$

How large of a sample, n , do we need for the approximation to be good?

Rule of Thumb: $np \geq 10$ and $n(1-p) \geq 10$

CLT Approximation to the Binomial When p is Close to 0.5

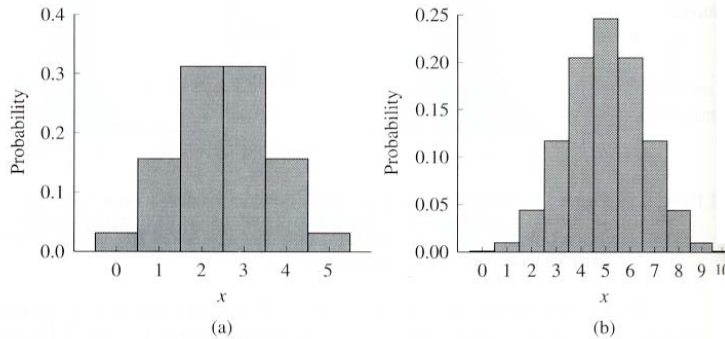


Figure 5.2 Probability Histograms of the Binomial distribution: (a) Binomial distribution with $p = 0.5$, $n = 5$; (b) Binomial distribution with $p = 0.5$, $n = 10$.

For a good approximation $np = n(1-p) = n0.5$ should be at least 10.
So, for a good approximation n should be at least 20

CLT Approximation to the Binomial When p is Not Close to 0.5

$np = n(.1)$
should be at least 10.
So n should be at least
100

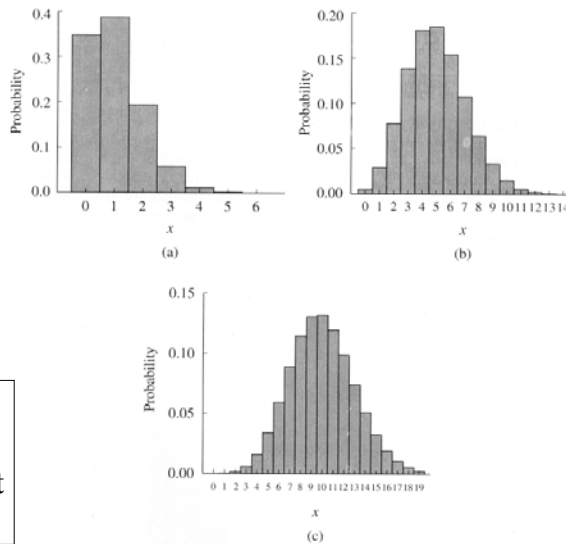


Figure 5.3 Probability Histograms of the Binomial Distribution: (a) Binomial distribution with $p = 0.1$, $n = 10$; (b) Binomial distribution with $p = 0.1$, $n = 50$; (c) Binomial distribution with $p = 0.1$, $n = 100$.

Continuity Correction

$$P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

$$P(X \leq 8) \approx \Phi\left(\frac{8.5 - np}{\sqrt{np(1-p)}}\right)$$

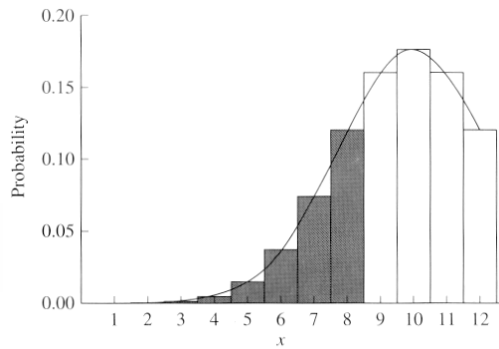


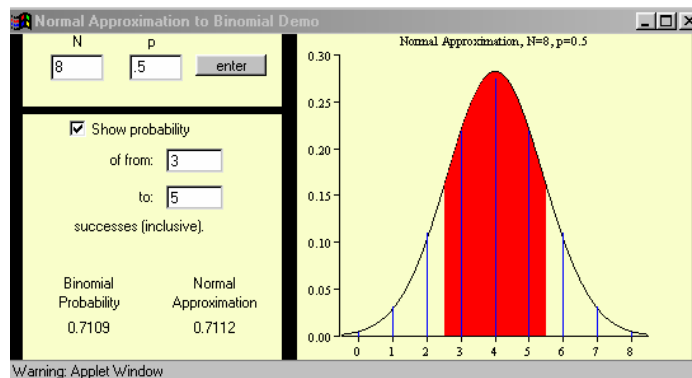
Figure 5.4 Normal Approximation to the Binomial Distribution

Similarly:

$$P(X \geq x) \approx 1 - \Phi\left(\frac{x - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

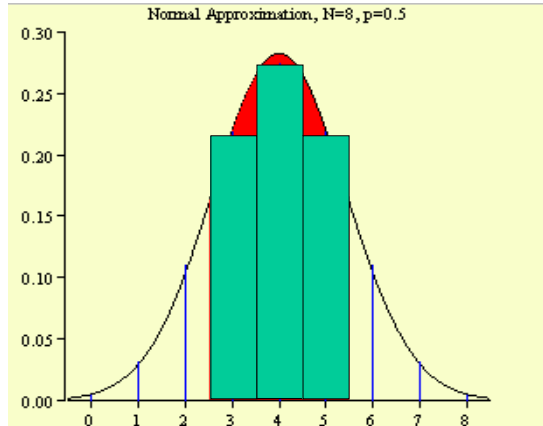
$$P(X \geq 8) \approx 1 - \Phi\left(\frac{7.5 - np}{\sqrt{np(1-p)}}\right)$$

Screen Shots of the Output of the Java Applet “Normal Approximation to the Binomial Distribution”



Homework: See the **Homework Log**.

Why the Normal Approximation to the Binomial Distribution Works in Pictures



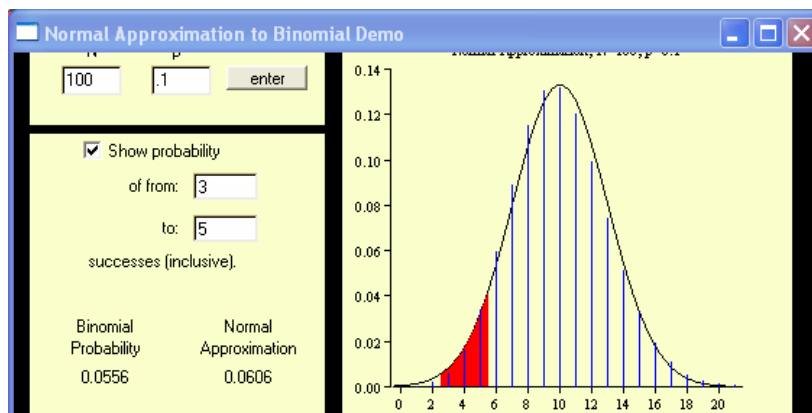
Green area is approximately the same as the red area

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

21

Java Applet for $N=100$ and $p=.1$



6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

22

Example: CLT Approximation to the Binomial

In a noisy communication channel there is a 2% chance that each transmitted bit (0 or 1) will be corrupted. If 1000 bits are transmitted, what is the probability that no more than 10 will be corrupted?

Let X = the number of corrupted bits. Then X has a binomial distribution with $n = 1000$ and $p = 0.02$. The required probability is

$$P(X \leq 10) = \sum_{x=0}^{10} \binom{1000}{x} (0.02)^x (0.98)^{1000-x}.$$

This probability is tedious to compute by hand. Using a computer package, we obtain the answer as 0.0102. The normal approximation can be computed more easily. We regard X as approximately normally distributed with $\mu = np = (1000)(0.02) = 20$ and $\sigma = \sqrt{np(1-p)} = \sqrt{(1000)(0.02)(0.98)} = 4.427$. Therefore

$$P(X \leq 10) = P\left(Z = \frac{X - \mu}{\sigma} \leq \frac{10 - 20}{4.427}\right) \approx \Phi(-2.259) = 0.0118.$$

This approximate probability is quite close to the exact binomial probability. ♦

Rolling Two Dice

Each one of these 36 outcomes are equally likely, i.e., each one occurs with 1/36 probability.

Now we pay attention to the sample variance.

Table 5.1 All Possible Samples of Size 2 from a Uniform Distribution on 1, 2, ..., 6 and Associated Sample Means and Variances

(x_1, x_2)	\bar{x}	s^2	(x_1, x_2)	\bar{x}	s^2
(1, 1)	1.0	0.0	(4, 1)	2.5	4.5
(1, 2)	1.5	0.5	(4, 2)	3.0	2.0
(1, 3)	2.0	2.0	(4, 3)	3.5	0.5
(1, 4)	2.5	4.5	(4, 4)	4.0	0.0
(1, 5)	3.0	8.0	(4, 5)	4.5	0.5
(1, 6)	3.5	12.5	(4, 6)	5.0	2.0
(2, 1)	1.5	0.5	(5, 1)	3.0	8.0
(2, 2)	2.0	0.0	(5, 2)	3.5	4.5
(2, 3)	2.5	0.5	(5, 3)	4.0	2.0
(2, 4)	3.0	2.0	(5, 4)	4.5	0.5
(2, 5)	3.5	4.5	(5, 5)	5.0	0.0
(2, 6)	4.0	8.0	(5, 6)	5.5	0.5
(3, 1)	2.0	2.0	(6, 1)	3.5	12.5
(3, 2)	2.5	0.5	(6, 2)	4.0	8.0
(3, 3)	3.0	0.0	(6, 3)	4.5	4.5
(3, 4)	3.5	0.5	(6, 4)	5.0	2.0
(3, 5)	4.0	2.0	(6, 5)	5.5	0.5
(3, 6)	4.5	4.5	(6, 6)	6.0	0.0

Sampling Distribution of the Sample Variance: Two Dice Example

Table 5.3 Sampling Distribution of S^2 for Samples of Size 2 from a Uniform Distribution on 1, 2, . . . , 6

s^2	0.0	0.5	2.0	4.5	8.0	12.5
$f(s^2)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

EXAMPLE 5.4 (ROLLING TWO DICE: SAMPLING DISTRIBUTION OF S^2)

Using the values of s^2 for all possible samples in Table 5.1, we obtain the distribution given in Table 5.3. Note that the distribution of S^2 is skewed to the right, in contrast to the distribution of \bar{X} given in Table 5.2, which is symmetric. This is quite typical; in general, the distribution of \bar{X} is approximately symmetric for modestly large n even if the population distribution is not symmetric (and exactly symmetric for any n if the population distribution is symmetric), while the distribution of S^2 is always skewed to the right. The mean of the distribution of S^2 is

$$E(S^2) = \frac{6}{36}(0) + \frac{10}{36}(0.5) + \frac{8}{36}(2.0) + \frac{6}{36}(4.5) + \frac{4}{36}(8.0) + \frac{2}{36}(12.5) = 2.9166,$$

which verifies that $E(S^2) = \sigma^2$. ◆

Chi-Square Distribution

Chi-Square Distribution. For $\nu \geq 1$, let Z_1, Z_2, \dots, Z_ν be i.i.d. $N(0, 1)$ r.v.'s and let

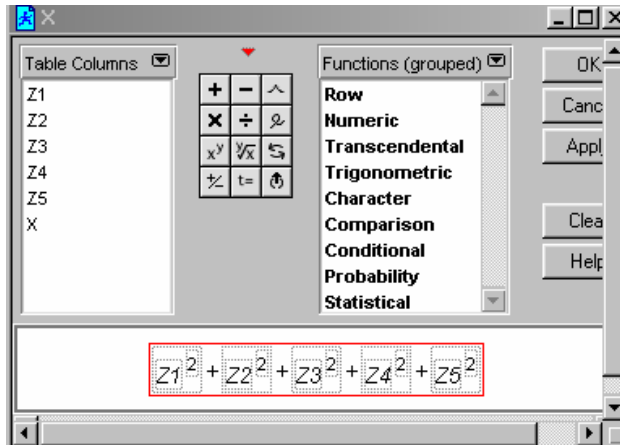
$$X = Z_1^2 + Z_2^2 + \dots + Z_\nu^2. \quad (5.3)$$

Then the probability density function (p.d.f.) of X can be shown to be

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} \quad \text{for } x \geq 0. \quad (5.4)$$

This is known as the χ^2 -distribution with ν degrees of freedom (d.f.), abbreviated as $X \sim \chi_\nu^2$. Note that (5.4) is a special case of the gamma distribution (see Section 2.8.2) with $\lambda = 1/2$ and $r = \nu/2$.

Using JMP to Simulate a Chi-Square Random Sample with 5 d.f.



The number of rows is the size of the random sample

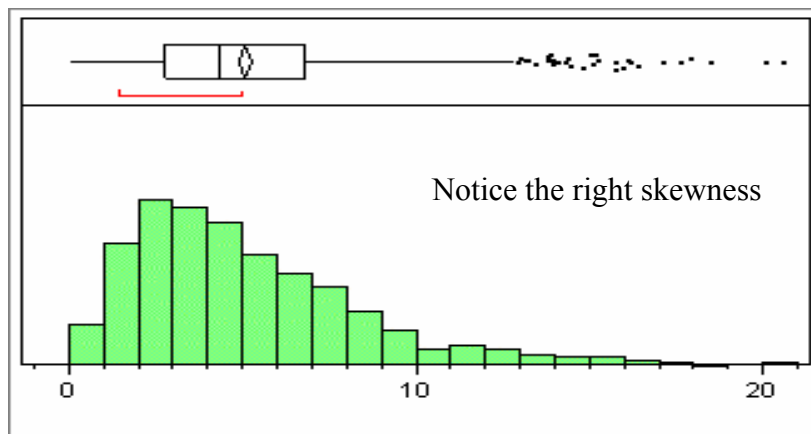
See the JMP tutorial “Chi-Square Simulation” on the course home page

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

27

Sample of 1000 Random Chi-Square Random Variables

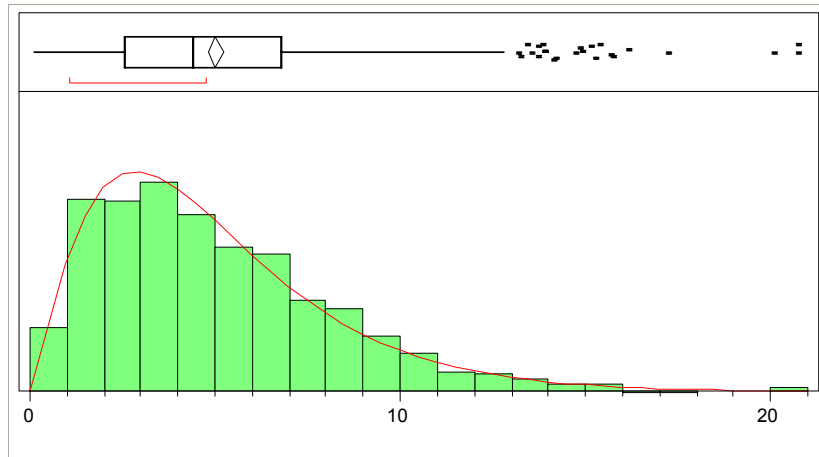


6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

28

Fitted Chi-Square Based on the Sample



6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

29

Chi-Square Density Function Curves

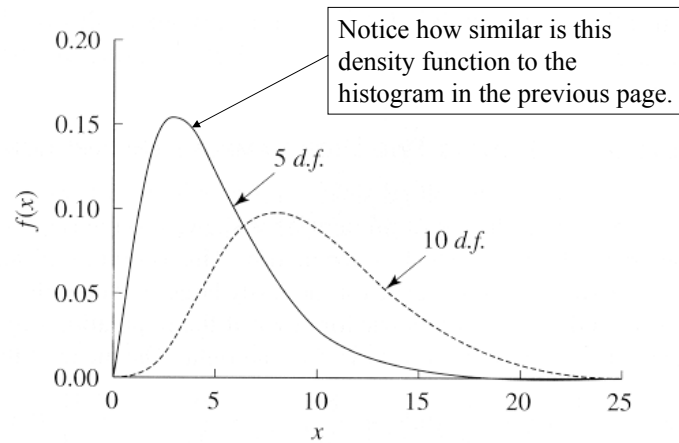


Figure 5.5 Chi-Square Density Function Curves for $\nu = 5$ and 10

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

30

Critical Values for the Chi-Square

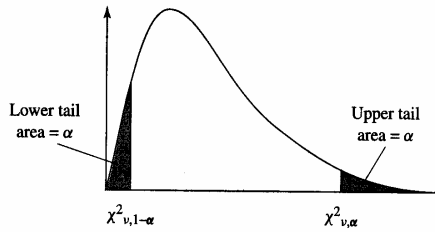
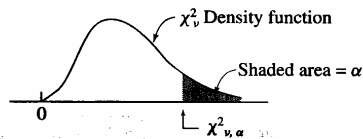


Figure 5.6 Lower and Upper α Critical Points of a Chi-Square Distribution

Table A.5 Critical Values $\chi^2_{v,\alpha}$ for the Chi-square Distribution



See the JMP tutorial "Tabled Values of Common Distributions"

v	α									
	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860

Distribution of Sample Variance

Assuming that the random sample comes from a normal distribution

It can be shown that (see the derivation given at the end of this section)

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \quad (5.7)$$

or equivalently

$$S^2 \sim \frac{\sigma^2 \chi^2_{n-1}}{n-1} \quad (5.8)$$

From the expressions for the mean and variance of the gamma distribution given in Section 2.8.2, we obtain

$$E(\chi^2_v) = v$$

and

$$\text{Var}(\chi^2_v) = 2v.$$

From this it follows that

$$E(S^2) = E\left(\frac{\sigma^2 \chi^2_{n-1}}{n-1}\right) = \frac{\sigma^2}{n-1} E(\chi^2_{n-1}) = \frac{\sigma^2}{n-1} \times (n-1) = \sigma^2 \quad (5.9)$$

Application of the Distribution of Sample Variance – Measurement Precision

Introduction to the ideas of hypothesis testing

The precision of an instrument refers to its ability to make repeat measurements that closely agree. One measure of precision is the variance among the repeat measurements; the smaller the variance the greater the precision. Suppose that an advertised claim for a thermometer is $\sigma = .01^\circ\text{F}$, but it is suspected that the true $\sigma > .01^\circ\text{F}$. To verify this claim, 10 repeat measurements are taken on a controlled temperature bath using this thermometer. If the observed sample variance s^2 is significantly larger than the claimed variance $\sigma^2 = (.01)^2$, this casts doubt on the advertised claim. In other words, if $s^2 > c(.01)^2$ for sufficiently large $c > 0$, then it is unlikely that $\sigma^2 = (.01)^2$. How large a value of c should be required to confidently reject the claim that the true $\sigma^2 = (.01)^2$?

Application of the Distribution of Sample Variance – Measurement Precision

One commonly used approach is to determine c so that there is only a 5% chance that S^2 will be larger than $c(.01)^2$ when in fact $\sigma^2 = (.01)^2$. Using this approach there is only a 5% chance of incorrectly concluding that the claim is unjustified, so we can be “95% confident” (this phrase is defined rigorously in the next chapter) in our rejection of the claim. Thus we need to find a value of c that satisfies

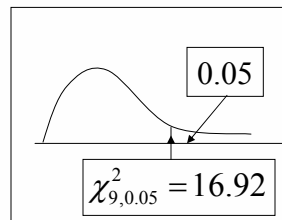
$$P\{S^2 > c(.01)^2\} = 0.05$$

when the claim $\sigma^2 = (.01)^2$ is true. We use the relationship

$$\frac{(n-1)S^2}{\sigma^2} = \frac{9S^2}{(.01)^2} \sim \chi_9^2$$

to write the above equation as

$$P\left\{\frac{9S^2}{(.01)^2} > \frac{9c(.01)^2}{(.01)^2}\right\} = P\{\chi_9^2 > 9c\} = 0.05.$$



Therefore $9c = \chi_{9,0.05}^2 = 16.92$ or $c = \frac{16.92}{9} = 1.88$. Thus, if s^2 exceeds $1.88(.01)^2$, we decide that the advertised claim is not supported by the data. ♦

Student's t -Distribution

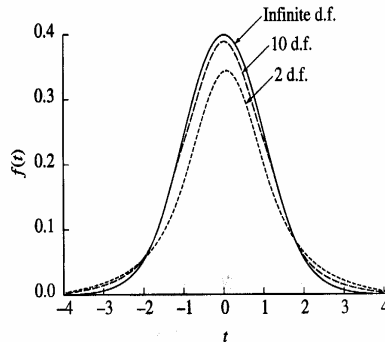
Consider a random sample, X_1, X_2, \dots, X_n drawn from a $N(\mu, \sigma^2)$

It is known that $\frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}}$ is exactly distributed as $N(0,1)$ for any n .

But $T = \frac{(\bar{X} - \mu)}{S/\sqrt{n}}$ is not longer distributed as $N(0,1)$.

The distribution of T is named Student's t -distribution.
(A different distribution for each number $\nu = n - 1 =$ degrees of freedom)

Play with the Java applet
"Student's t Distribution"



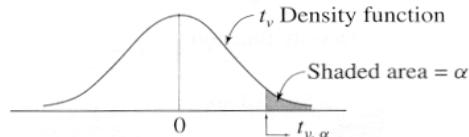
6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

35

t -Distribution Table

Table A.4 Critical Values $t_{\nu, \alpha}$ for the t -Distribution



ν	α						
	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
29	1.311	1.699	2.045	1.462	2.756	3.396	3.659
30	1.310	1.697	2.042	1.457	2.750	3.385	3.646
40	1.303	1.684	2.021	1.423	2.704	3.307	3.551
60	1.296	1.671	2.000	1.390	2.660	3.232	3.460
120	1.289	1.658	1.980	1.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	1.326	2.576	3.090	3.291

See the JMP tutorial "Tabled Values of Common Distributions"

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

36

Application of the t -Distribution Calculation – Process Control

Half-pint (8 oz.) milk cartons are filled at a dairy by a filling machine. To provide a check on the machine, a sample of 10 cartons is periodically measured. If the sample mean \bar{x} deviates by more than a certain amount d from the nominal value of 8 oz., i.e., if $|\bar{x} - 8| > d$, then the machine setting is adjusted. The chance of a false alarm indicating an unnecessary adjustment is to be limited to 1%. Find a formula for d .

If the true mean amount of milk, denoted by μ , equals 8 oz. then no adjustment is required. Therefore the probability of a false alarm is given by

$$P(|\bar{X} - 8| > d | \mu = 8) = 0.01.$$

Example: t -Distribution Calculation

If the true mean amount of milk, denoted by μ , equals 8 oz. then no adjustment is required. Therefore the probability of a false alarm is given by

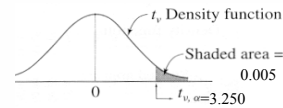
$$P(|\bar{X} - 8| > d | \mu = 8) = 0.01.$$

Divide both sides of the inequality in the probability statement by S/\sqrt{n} , where $n = 10$, and use the fact that $\frac{\bar{X}-8}{S/\sqrt{10}}$ has a t -distribution with $10 - 1 = 9$ d.f. when $\mu = 8$. Therefore the above equation becomes

$$P\left(|T_9| > \frac{d}{S/\sqrt{10}}\right) = 0.01.$$

Let

$$c = \frac{d}{S/\sqrt{10}} \quad \text{or} \quad d = \frac{cS}{\sqrt{10}}.$$



Then c must cut off an area $0.01/2 = 0.005$ in each tail of the t -distribution with 9 d.f., i.e., $c = t_{9, 0.005} = 3.250$ from Table A.4. Thus the decision rule decides that the machine needs adjustment if

$$|\bar{x} - 8| > d = 3.250 \frac{s}{\sqrt{10}}. \quad \blacklozenge$$

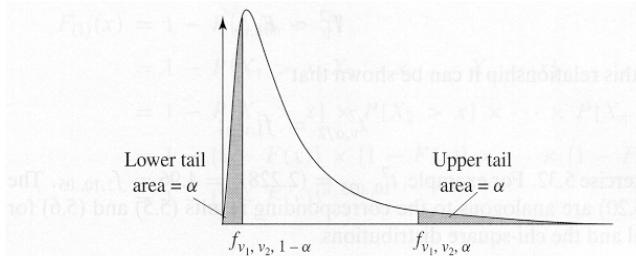
F-Distribution

Consider two independent random samples,

X_1, X_2, \dots, X_{n_1} from an $N(\mu_1, \sigma_1^2)$, Y_1, Y_2, \dots, Y_{n_2} from an $N(\mu_2, \sigma_2^2)$.

Then $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ has an *F distribution* with $v_1 = n_1 - 1$ d.f.

in the numerator and $v_2 = n_2 - 1$ d.f. in the denominator.



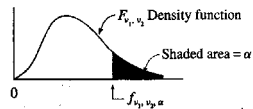
6/12/2004

Figure 5.8 Lower and Upper α Critical Points of an *F*-Distribution

39

F-Distribution Table

Table A.6 Critical Values $f_{v_1, v_2, \alpha}$ for the *F*-Distribution ($\alpha = .05$) (cont.)



Degrees of freedom for the denominator (v_2)	Degrees of freedom for the numerator (v_1)																			
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40	
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.49	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	
17	4.45	3.59	3.20	2.96	2.81	2.69	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76	
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73	

See the JMP tutorial "Tabled Values of Common Distributions"

6/12/2004

Unit 5 - Stat 571 - Ramon V. Leon

40