

EEC 686/785

Modeling & Performance Evaluation of Computer Systems

Lecture 8

Wenbing Zhao

Department of Electrical and Computer Engineering
Cleveland State University

wenbing@ieee.org

(based on Dr. Raj jain's lecture notes)

Outline

- Review of lecture 7
- Comparing Systems Using Sample Data
- **Reminder: Midterm #1 next Monday Oct. 3, 2005**

Basic Probability and Statistics Concepts

- **Independent events**
- **Random variables**
- **Cumulative Distributed Function (CDF):** $F_x(a) = P(x \leq a)$
- **Probability Density Function:** $f(x) = dF(x)/dx$
- **Probability Mass Function:** $f(x_i) = p_i$
- **Mean or Expected Value** $\mu = E(x) = \sum_{i=1}^n p_i x_i = \int_{-\infty}^{+\infty} x f(x) dx$
- **Variance (σ^2):** $Var(x) = E[(x - \mu)^2]$
- **Standard Deviation:** σ

Basic Probability and Statistics Concepts

- **Coefficient of Variation (C.O.V.):** σ/μ
- **Covariance:** $Cov(x,y) = \sigma_{xy}^2 = E[(x-\mu_x)(y-\mu_y)] = E(xy) - E(x)E(y)$
- **Correlation Coefficient:** $\rho_{xy} = \sigma_{xy}^2 / \sigma_x \sigma_y$
- **Quantile:** $P(x \leq x_\alpha) = F(x_\alpha) = \alpha$
- **Medium:** the 50-percentile
- **Mode**
- **Normal distribution:** $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, -\infty \leq x \leq +\infty$

Selecting among the Mean, Median and Mode ⁵

4 October 2005

EEEC686/785

Wenbing Zhao

Geometric Mean ⁶

- The **geometric mean** of n values x_1, x_2, \dots, x_n is obtained by multiplying the values together and taking the n th root of the product:

$$\dot{x} = \left(\prod_{i=1}^n x_i \right)^{1/n}$$

- The **arithmetic mean** is used if the *sum of the observations* is a quantity that is of interest
- The **geometric mean** is used if the *product of the observations* is a quantity of interest
 - Average error rate per hop on a multihop path in a network

4 October 2005

EEEC686/785

Wenbing Zhao

Harmonic Mean ⁷

- The **harmonic mean** of a sample $\{x_1, x_2, \dots, x_n\}$ is defined as follows:

$$\ddot{x} = \frac{n}{1/x_1 + 1/x_2 + \dots + 1/x_n}$$

- A harmonic mean should be used whenever an arithmetic mean can be justified for $1/x_i$

4 October 2005

EEEC686/785

Wenbing Zhao

Rules for Summarizing the Ratios ⁸

- If we take the sum of numerators and the sum of denominators and both have a physical meaning, the average of the ratio is the ratio of the averages
- When the arithmetic mean of the ratios can be used?
 - Denominator is a constant, so that the ratio has been taken with respect to a base that is constant across all observations,
 - The sum of the numerator has a physical meaning
- When a harmonic mean of the ratio can be used
 - The sum of the denominators has a physical meaning
 - The numerators are constant
- When a geometric mean can be used
 - The numerator and the denominator are expected to follow a multiplicative property such that $a_i = cb_i$, where c is approximately a constant that is being estimated, =>
 - c can be estimated by the geometric mean of a_i/b_i

4 October 2005

EEEC686/785

Wenbing Zhao

Indices of Dispersion

- **Indices of dispersion:** variability is specified using one of the following measures
 - Range – minimum and maximum of the values observed
 - Variance or standard deviation
 - 10- and 90- percentiles
 - Semi-interquartile range (SIQR)
 - Mean absolute deviation

Selecting the Index of Dispersion

Determining Distribution of Data

- Sometimes it is useful and necessary to state the **type of distribution** of the data follows, in addition to its **average and variability**
 - E.g., number of disk I/O's are uniformly distributed between 1 and 25 => more meaning than to say the mean is 13 and the variance of 48
 - The distribution information is also required if the summary has to be used later in simulation or analytical modeling

Determining Distribution of Data

- A simple way to determine the distribution is to plot a histogram of the observations
 - Determining the max and min of the values observed
 - Dividing the range into a number of subranges called cells or buckets
 - The count of observations that fall into each cell is determined
 - The counts are normalized to cell frequencies by dividing by the total number of observations
 - The cell frequencies are plotted as a column chart

Determining Distribution of Data

- The key problem in plotting histograms is determining the cell size
- One guideline is that if any cell has less than five observations, the cell size should be increased or a variable cell histogram should be used
- A better technique for small samples is to plot the **observed quantiles** versus the **theoretical quantile** in a **quantile-quantile** plot
 - If the observations do come from the given theoretical distribution, the quantile-quantile plot would be linear

Determining Distribution of Data

- Quantile-quantile plot:
 - Suppose, $y_{(i)}$ is the observed q_i^{th} quantile
 - Using the theoretical distribution, the q_i^{th} quantile x_i is computed and a point is plotted at $(x_i, y_{(i)})$
- To determine the q_i^{th} quantile x_i , we need to invert the cumulative distribution function. If $F(x)$ is the CDF for the assumed distribution,

$$q_i = F(x_i) \quad \text{or} \quad x_i = F^{-1}(q_i)$$

- For the unit normal distribution $N(0,1)$, the following approximation is often used:

$$x_i = 4.91[q_i^{0.14} - (1 - q_i)^{0.14}]$$

Example 12.5

Interpretation of Normal Quantile-Quantile Plots

- If the data follows a straight line but departs from it at one or both ends, this indicates that the data has shorter or longer tails than the normal distribution

Comparing Systems Using Sample Data

- Sample
 - Old French word 'essample' => 'sample' and 'example'
 - One example \neq theory
 - One sample \neq definite statement of the system

Sample versus Population

- Generate several million random numbers with mean μ and standard deviation σ
- Draw a sample of n observations
sample mean $\bar{x} \neq$ population mean μ
- **Parameters:** population characteristics = unknown = Greek
statistics: sample estimates = random = English

Confidence Interval for The Mean

- k samples => k sample means => can't get a perfect estimate of μ from any finite number of finite size samples => use probabilistic bounds c_1 and c_2 :

$$\text{Probability } \{c_1 \leq \mu \leq c_2\} = 1-\alpha$$

- **Confidence interval** for the population mean: (c_1, c_2)
- **Significance level:** α
- **Confidence level:** $100(1-\alpha)$
- **Confidence coefficient:** $1-\alpha$

Determine Confidence Interval

- Use 5-percential and 95-percentile of the sample means to get 90 => need many samples
- **Central limit theorem:** if the observations in a sample are independent and come from the same population that has a mean μ and a standard deviation σ , then, the sample mean for large samples is approximately normally distributed with mean μ and standard deviation σ/\sqrt{n}

$$\bar{x} \sim N(\mu, \sigma/\sqrt{n})$$

- **Standard error:** standard deviation of the sample mean =

$$\sigma/\sqrt{n}$$

Determine Confidence Interval

- 100(1- α)% confidence interval for the population mean:

$$(\bar{x} - z_{1-\alpha/2}s/\sqrt{n}, \bar{x} + z_{1-\alpha/2}s/\sqrt{n})$$

$$z_{1-\alpha/2} = (1 - \alpha/2)\text{-quantile of } N(0,1)$$

Example 13.1

- For the sample of example 12.4
 - $\bar{x} = 3.90, s=0.95$ and $n=32$
- A 90% confidence interval for the mean = $3.90 \mp (1.645)(0.95)/\sqrt{32} = (3.62, 4.17)$
- We can state with 90% confidence that the population mean is between 3.62 and 4.17
- The chance of error in this statement is 10%
- A 95% confidence interval for the mean = $3.90 \mp (1.960)(0.95)/\sqrt{32} = (3.57, 4.23)$
- A 99% confidence interval for the mean = $3.90 \mp (2.576)(0.95)/\sqrt{32} = (3.46, 4.33)$

Confidence Interval: Meaning

- 90% confidence:
 - If we take 100 samples and construct confidence interval for each sample, the interval would include the population mean in 90 cases

Confidence Interval for Small Samples

- 100(1- α)% confidence interval for $n < 30$:

$$(\bar{x} - t_{[1-\alpha/2; n-1]}s/\sqrt{n}, \bar{x} + t_{[1-\alpha/2; n-1]}s/\sqrt{n})$$

$t_{[1-\alpha/2; n-1]} = (1 - \alpha/2)$ -quantile of a t -variate with $n - 1$ degree of freedom

$$x \sim N(\mu, \sigma^2) \Rightarrow (\bar{x} - \mu)/(\sigma/\sqrt{n}) \sim N(0,1)$$

$$(n-1)s^2/\sigma^2 \sim \chi^2(n-1) \Rightarrow (\bar{x} - \mu)/\sqrt{s^2/n} \sim t(n-1)$$

Sample t Density Function

- The value $t_{[1-\alpha/2;n-1]}$ is such that the
 - Probability of the random variable being less than $-t_{[1-\alpha/2;n-1]}$ is $\alpha/2$
 - Probability of the random variable being more than $t_{[1-\alpha/2;n-1]}$ is $\alpha/2$
 - The probability that the variable will be between $\pm t_{[1-\alpha/2;n-1]}$ is $1-\alpha$

Example

- Error data of example 12.5:
-0.04, -0.19, 0.14, -0.09, -0.14, 0.19, 0.04, and 0.09
- Mean = 0, sample standard deviation = 0.138
- The $t_{[0.95;7]}$ from Table A.4 is 1.895, thus,
- The 90% confidence interval for the mean:
 $0 \mp (1.895)(0.138) / \sqrt{8} = 0 \mp 0.0926 = (-0.0926, 0.0926)$

Testing for a Zero Mean

- The test consists of determining a confidence interval and simply checking if the interval includes zero
- The procedure for testing for zero mean applies equally well to any other value as well

Example 13.3

- The measured difference in the processor times of two different implementations of the same algorithm:
{1.5, 2.6, -1.8, 1.3, -0.5, 1.7, 2.4}
- Can we say with 99% confidence that one is superior to the other?
 - Sample size $n = 7$; mean = $7.20/7 = 1.03$
 - Sample variance = $(22.84 - 7.20 * 7.20/7)/6 = 2.57$
 - Sample standard deviation = $1.60 (\sqrt{2.57})$
 - Confidence interval = $1.03 \pm t * 1.60 / \sqrt{7} = 1.03 \pm 0.605t$
 - $100(1-\alpha) = 99$, $\alpha = 0.01$, $1-\alpha/2 = 0.995 \Rightarrow t_{[0.995;6]} = 3.707$
 - 99% confidence interval = $(-1.21, 3.27)$
- We cannot say with 99% confidence that the mean difference is significantly different from 0 \Rightarrow they are the same

Comparing Two Alternatives: Paired vs. Unpaired

29

- Paired: one-to-one correspondence between the i th test on system B
 - Example: performance on i th workload
=> use confidence interval of the difference
- Unpaired: no correspondence
 - Example: n people on system A, n on system B
=> need more sophisticated method

4 October 2005

EEEC686/785

Wenbing Zhao

Paired Observations: Example

30

- Performance: $\{(5.4, 19.1), (16.6, 3.5), (0.6, 3.4), (1.4, 2.5), (0.6, 3.6), (7.3, 1.7)\}$
- Is one system better than the other?
 - Differences: $\{-13.7, 13.1, -2.8, -1.1, -3.0, 5.6\}$
 - Sample mean = -0.32
 - Sample variance = 81.62
 - Sample standard deviation = 9.03
 - Confidence interval for the mean = $-0.32 \pm t \sqrt{(81.62/6)} = -0.32 \pm t(3.69)$
 - $t_{[0.95,5]} = 2.015$
 - 90% confidence interval = $-0.32 \pm (2.015)(3.69) = (-7.75, 7.11)$
- Answer: No. They are not different

4 October 2005

EEEC686/785

Wenbing Zhao

Unpaired Observations

31

- t -test procedure
 1. Compute the sample means:

$$\bar{x}_a = \frac{1}{n_a} \sum_{i=1}^{n_a} x_{ia} \quad \bar{x}_b = \frac{1}{n_b} \sum_{i=1}^{n_b} x_{ib}$$

2. Compute the sample standard deviations

$$s_a = \left\{ \frac{(\sum_{i=1}^{n_a} x_{ia}^2) - n_a \bar{x}_a^2}{n_a - 1} \right\}^{\frac{1}{2}} \quad s_b = \left\{ \frac{(\sum_{i=1}^{n_b} x_{ib}^2) - n_b \bar{x}_b^2}{n_b - 1} \right\}^{\frac{1}{2}}$$

4 October 2005

EEEC686/785

Wenbing Zhao

Unpaired Observations

32

- t -test procedure
 3. Compute the mean difference: $(\bar{x}_a - \bar{x}_b)$
 4. Compute the standard deviation of the mean difference:

$$s = \sqrt{\left(\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b} \right)}$$

4 October 2005

EEEC686/785

Wenbing Zhao

Unpaired Observations

- *t*-test procedure

5. Compute the effective number of degrees of freedom:

$$\nu = \frac{\left(\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}\right)^2}{\frac{1}{n_a+1}\left(\frac{s_a^2}{n_a}\right)^2 + \frac{1}{n_b+1}\left(\frac{s_b^2}{n_b}\right)^2} - 2$$

6. Compute the confidence interval for the mean difference: $(\bar{x}_a - \bar{x}_b) \mp t_{[1-\alpha/2, \nu]} s$

Unpaired Observations: Example

- Times on system A: {5.36, 16.57, 0.62, 1.41, 0.64, 7.26}
- Times on system B: {19.12, 3.52, 3.38, 2.50, 3.60, 1.76}

- Question: are the two systems significantly different?

- For system A: mean $\bar{x}_a = 5.31$; variance $s_a^2 = 37.92$; $n_a = 6$
- For system B: mean $\bar{x}_b = 5.64$; variance $s_b^2 = 44.11$; $n_b = 6$
- Mean difference $\bar{x}_a - \bar{x}_b = -0.33$
- Standard deviation of the mean difference = 3.698
- Effective number of degree of freedom $f=11.921$
- The 0.95-quantile of a *t*-variate with 12 degrees of freedom is 1.71
- The 90% confidence interval for the difference = (-0.692, 6.26)

- The confidence interval includes zero => two systems are not different

Approximate Visual Test

- To compare two unpaired samples, simply compute the confidence interval for each alternative separately

Approximate Visual Test

- Three possibilities:

- Confidence intervals do not overlap. The one with higher sample mean is significantly better
- Confidence intervals overlap considerable such that the mean of one falls in the interval for the other => The two are equal with the desired confidence
- Confidence intervals overlap slightly such that the mean of either is outside the confidence interval for the other => We need to do the *t*-test

Approximate Visual Test: Example

- Times on system A: {5.36, 16.57, 0.62, 1.41, 0.64, 7.26}
Times on system B: {19.12, 3.52, 3.38, 2.50, 3.60, 1.76}
- $t_{[0.95,5]}=2.015$
- The 90% confidence interval for the mean of A = $5.31 \pm (2.015)\sqrt{(37.92/6)}$
= (0.24,10.38)
- The 90% confidence interval for the mean of B = $5.64 \pm (2.015)\sqrt{(44.11/6)}$
= (0.18,11.10)
- Confidence intervals overlap and the mean of one falls in the confidence interval for the other
=> two systems are not different at this level confidence

What Confidence Level To Use?

- Need not always be 90% or 95% or 99%
- Base on the loss that you would sustain if the parameter is outside the range and the gain you would have if the parameter is inside the range
- Low loss => Low confidence level is find
 - > E.g., lottery \$5 million with probability 10^{-7}
 - > 90% confidence => buy 9 million tickets (no one would be willing)
 - > 0.01% confidence level is fine with most people
- 50% confidence level may or may not be too low
- 99% confidence level may or may not be too high

Hypothesis Testing Versus Confidence Intervals

- Confidence Interval (CI) or simple hypothesis testing
- Confidence interval provides more information
- Hypothesis test = yes-no decision
- Confidence interval also provides possible range
- Narrow confidence interval => high degree of precision
- Wide confidence interval => low precision

Hypothesis Testing Versus Confidence Intervals

- Example:
 - > (-100, 100) => no difference
 - > (-1, 1) => no difference
- Confidence intervals tell us not only what to say but also how loudly to say
- CI is easier to explain to decision makers
- CI is more useful, for example
 - > Parameter range (100, 200) vs. probability of (parameter = 100) = 3%

One Sided Confidence Intervals

- In all the tests so far, two-sided intervals have been used
 - For such intervals, if the confidence level is $100(1-\alpha)\%$
 - \Rightarrow there is a $100\alpha/2\%$ chance the difference will be more than the upper confidence limit, and
 - \Rightarrow a $100\alpha/2\%$ chance the difference will be less than the lower confidence limit
 - E.g., 90% confidence
 - $\Rightarrow P(\text{difference} > \text{upper limit}) = 5\%$
 - $\Rightarrow P(\text{difference} < \text{lower limit}) = 5\%$

One Sided Confidence Intervals

- One sided question: is the mean greater than 0?
 \Rightarrow One side confidence interval
- One sided lower confidence interval for μ :

$$\left(\bar{x} - t_{[1-\alpha; n-1]} \frac{s}{\sqrt{n}}, \infty\right)$$

Note t at $1-\alpha$ (not $1-\alpha/2$)
- One sided upper confidence interval for μ :

$$\left(-\infty, \bar{x} + t_{[1-\alpha; n-1]} \frac{s}{\sqrt{n}}\right)$$
- For large samples: use z instead of t

One Sided Confidence Intervals: Example

- Time between crashes

System	Number	Mean	Stdv
A	972	124.10	198.20
B	153	141.47	226.11

- Assume unpaired observations
- Mean difference:

$$\bar{x}_A - \bar{x}_B = 124.10 - 141.47 = -17.37$$
- Standard deviation of the difference:

$$s = \sqrt{\left(\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}\right)} = \sqrt{\frac{(198.20)^2}{972} + \frac{(226.11)^2}{153}} = 19.35$$

One Sided Confidence Intervals: Example

- Effective number of degrees of freedom:

$$\begin{aligned} v &= \frac{\left(\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}\right)^2}{\frac{1}{n_a+1} \left(\frac{s_a^2}{n_a}\right)^2 + \frac{1}{n_b+1} \left(\frac{s_b^2}{n_b}\right)^2} - 2 \\ &= \frac{\left(\frac{(198.20)^2}{972} + \frac{(226.11)^2}{153}\right)^2}{\frac{1}{972+1} \left(\frac{(198.20)^2}{972}\right)^2 + \frac{1}{153+1} \left(\frac{(226.11)^2}{153}\right)^2} - 2 \\ &= 191.05 \end{aligned}$$

One Sided Confidence Intervals: Example

- $\nu > 30 \Rightarrow$ use z rather than t
- One sided test \Rightarrow use $z_{0.90} = 1.28$ for 90% confidence
- 90% confidence interval:
 $(-\infty, -17.37 + 1.28 \times 19.35) = (-\infty, 7.402)$
- CI includes zero \Rightarrow System A is not more susceptible to crashes than system B

Confidence Intervals for Proportions

- Proportion = probability of various categories
 - E.g., $P(\text{error})=0.01$, $P(\text{no error})=0.99$
- n_1 for n observations are of type 1 \Rightarrow
 - Sample proportion = $p = n_1/n$
 - Confidence interval for the proportion $p \mp z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}}$
- Assumes normal approximation of Binomial distribution \Rightarrow Valid only if $np \geq 10$
- Need to use binomial tables if $np < 10$, can't use t -values

Example 13.9

- 10 out of 1000 pages printed on a laser printer are illegible
 - Sample proportion = $p = 10/1000 = 0.01$
- $np \geq 10$, can apply the formula:
 - Confidence interval $= p \mp z \sqrt{\frac{p(1-p)}{n}}$
 $= 0.01 \mp z \sqrt{\frac{0.01(0.99)}{1000}} = 0.01 \mp 0.003z$
 - 90% confidential interval = $0.01 \pm (1.645)(0.003) = (0.005, 0.015)$
 - 95% confidential interval = $0.01 \pm (1.960)(0.003) = (0.004, 0.016)$
- At 90% confidence: 0.5% to 1.5% of the pages are illegible
 Chances of error of this statement = 10%
- At 95%, chance of error of a similar statement = 5%

Example 13.10

- 40 Repetitions on two systems: system A superior in 26 repetitions
- Question: with 99% confidence, is system A superior?
- $P = 26/40 = 0.65$
 Standard deviation = $\sqrt{p(1-p)/n} = 0.075$
 99% confidence interval = $0.65 \pm (2.576)(0.075) = (0.46, 0.84)$
- CI includes 0.5 \Rightarrow We cannot say with 99% confidence that system A is superior
- 90% does not include 0.5 \Rightarrow Can say with 90% confidence that system A is superior

Sample Size for Determining Mean

- Larger sample \Rightarrow Narrower confidence interval \Rightarrow Higher confidence
- Question: How many observations n to get an accuracy of $\pm r\%$ and a confidence level of $100(1-\alpha)\%$?
 - The $100(1-\alpha)\%$ confidence interval of the population mean is

$$\bar{x} \mp z \frac{s}{\sqrt{n}}$$

- $r\%$ accuracy $\Rightarrow CI = (\bar{x}(1-r/100), \bar{x}(1+r/100))$

$$\Rightarrow \bar{x} \mp z \frac{s}{\sqrt{n}} = \bar{x} \left(1 \mp \frac{r}{100} \right) \Rightarrow z \frac{s}{\sqrt{n}} = \bar{x} \frac{r}{100} \Rightarrow n = \left(\frac{100zs}{r\bar{x}} \right)^2$$

Example 13.11

- Sample mean of the response time = 20 seconds
Sample standard deviation = 5
- Question: How many repetitions are needed to get the response time accurate within 1 second at 95% confidence?

- Required accuracy = 1 in 20 = 5%
Here, $\bar{x} = 20, s = 5, z = 1.960$, and $r = 5$,

$$n = \left(\frac{(100)(1.960)(5)}{(5)(20)} \right)^2 = (9.8)^2 = 96.04$$

A total of 97 observations are needed

Sample Size for Determining Proportions

- Confidential interval for the proportion

$$p \mp z \sqrt{\frac{p(1-p)}{n}}$$

- To get a half-width (accuracy of) r : $p \mp r = p \mp z \sqrt{\frac{p(1-p)}{n}}$

$$r = z \sqrt{\frac{p(1-p)}{n}}$$

$$n = z^2 \frac{p(1-p)}{r^2}$$

Example 13.12

- Preliminary measurement: illegible print rate of 1 in 10,000
- Question: How many pages must be observed to get an accuracy of 1 per million at 95% confidence?
- Answer: $p = 1/10000 = 1E-4, r = 1E-6, z = 1.960$

$$n = (1.960)^2 \left(\frac{10^{-4}(1-10^{-4})}{(10^{-6})^2} \right) = 384160000$$

A total of 384.16 million pages must be observed

Sample Size for Comparing Two Alternatives: Example

53

- Algorithm A loses 0.5% of packets and algorithm B loses 0.6%
- Question: How many packets do we need to observe to state with 95% confidence that algorithm A is better than the algorithm B?

4 October 2005

EEC686/785

Wenbing Zhao

Sample Size for Comparing Two Alternatives: Example

54

Answer:
 CI for algorithm A = $0.005 \mp 1.960 \left(\frac{0.005(1-0.005)}{n} \right)^{1/2}$

CI for algorithm B = $0.006 \mp 1.960 \left(\frac{0.006(1-0.006)}{n} \right)^{1/2}$

- For non-overlapping intervals:

$$0.005 \mp 1.960 \left(\frac{0.005(1-0.005)}{n} \right)^{1/2} \leq 0.006 \mp 1.960 \left(\frac{0.006(1-0.006)}{n} \right)^{1/2}$$

$$n \geq 84340$$

We need to observe 85,000 packets

4 October 2005

EEC686/785

Wenbing Zhao

Summary

55

- n_1 of the n observations belong to a certain class, the following statistics can be reported for the class:
 - Proportion of the observations in the class: $p = n_1/n$
 - 100(1- α)% two sided confidence interval for the proportion (provided $np \geq 10$):

$$p \mp z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

- 100(1- α)% one sided confidence interval for the proportion (provided $np \geq 10$):

$$\left(p, p + z_{1-\alpha} \sqrt{\frac{p(1-p)}{n}} \right) \quad \text{or} \quad \left(p - z_{1-\alpha} \sqrt{\frac{p(1-p)}{n}}, p \right)$$

4 October 2005

EEC686/785

Wenbing Zhao