Chapter 4
Confidence Interval of the Mean

Section 4.3
The Confidence Interval

David W. Gerbing

The School of Business
Portland State University

- The Confidence Interval in Practice
  - Estimated Range of Variation
  - Construction, Meaning and Interpretation
  - Application

4.3a
Estimated Range of Variation
t-Distributions and Cutoff Values

t-value: Estimated range of variation

In practice, σ is unknown, so a price to pay

- **Result**: 95% of all sample means, m, are within $1.96\sigma_m$'s of μ
- **Problem**: With actual data analysis, σ is not known, and so then neither is the true standard error, $\sigma_m = \sigma/\sqrt{n}$
- When analyzing data, replace σ with its estimate from a *single* sample, the standard deviation of the sample data, s
- ***Estimated* (sample) standard error of m**: $s_m = \dfrac{s}{\sqrt{n}}$
- To assess the distance of the sample mean, m, from the true (known) mean, μ, with information s and n calculated from a single sample, define a *new* standardized value
- **t-value** or t-statistic: $t_m = \dfrac{m-\mu}{s_m}$ vs. $z_m = \dfrac{m-\mu}{\sigma_m}$
- **Key Concept**: A t-value is how many *estimated* standard errors separate m from μ, instead of the *actual* standard errors

The Concept of a t-Distribution

A mathematical construction of what *would* happen

- Can calculate a t-value from the values of n, m, and s from each of many samples, but best done with mathematics . . .
- **t-distribution**: Mathematically defined distribution of t-values over many, usually hypothetical, random samples, all of the same size n from the same population of data Y
- A distribution of t-values also plots as a bell-shaped curve
- However, unlike z_m in which only m varies across samples, for t_m, *both* m and s_m differ from sample to sample
- **Key Concept**: There is a different t-distribution for each sample size, n, in terms of degrees of freedom, $df = n - 1$
- The larger the sample size, n, the better s estimates σ . . .
  - The less s, and thus s_m, fluctuates from sample to sample
  - The narrower and less variable is the corresponding t-distribution, increasingly resembling normality

Estimated Range of Variation

Usually set the range of variation of sampling error at 95%

- **Key Concept**: The practical implication is that the estimate of the range of sampling variation is wider than the actual range
- For the 95% range of variation, how much larger is the relevant t-cutoff compared to the baseline z-cutoff of 1.96 from the standardized normal curve?
- Show the probability interval for the 95% range of variation for both the relevant t-distribution and the z-distribution, that is, the standardized normal curve

- Illustrate with the `lessR` function for t-cutoffs: `prob_tcut()`

Required: `df`, degrees of freedom, $n - 1$

Default: `alpha=0.05`, which sets a 95% confidence level

Optional: `?prob_tcut`, e.g., `col_fill="aliceblue"`

$z_{.025}$ vs. $t_{.025,df=5}$

```
> prob_tcut(df=5)
```

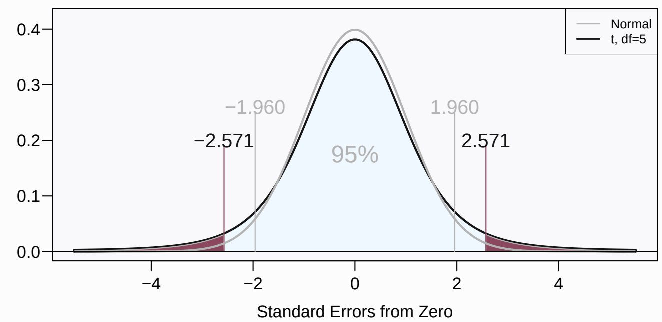


Figure: t-distribution, $df = 5$, $t_{.025} = 2.571$, and 95% probability interval

$z_{.025}$ vs. $t_{.025}$ for $df = 15$

```
> prob_tcut(df=15)
```

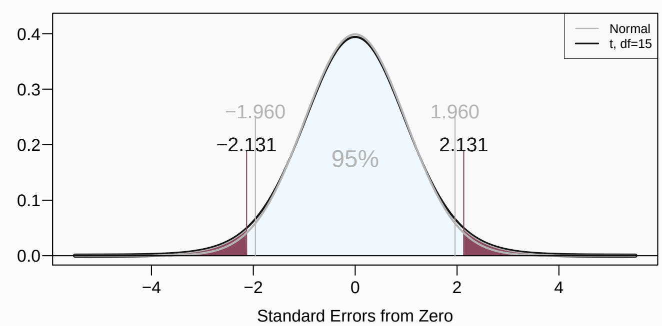


Figure: t-distribution, $df = 15$, $t_{.025} = 2.131$, and 95% probability interval

$z_{.025}$ vs. $t_{.025}$ for $df = 200$

```
> prob_tcut(df=200)
```

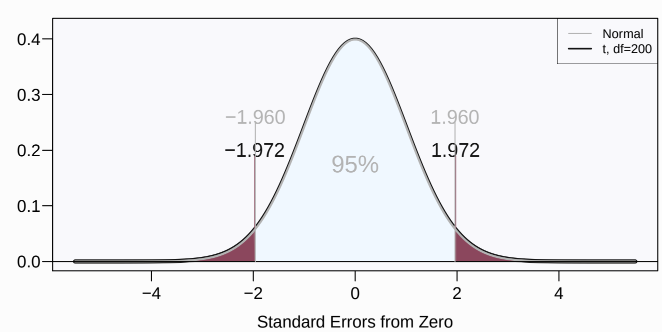


Figure: t-distribution, $df = 200$, $t_{.025} = 1.972$, and 95% probability interval

Summary of $t_{.025}$ cutoffs

95% range of sampling variation of t_m from sample size

- What range of t-values contains 95% of these values over the hypothetical many samples?
- **Notation**: $t_{.025}$ is the cutoff that defines 2.5% of the upper tail for the t-distribution at degrees of freedom, $df = n - 1$
- *Range of sampling variability*: 95% of t-values, t_m, between $-t_{.025}$ and $t_{.025}$
- *Most basic summary*: $t_{.025} \approx 2 > z_{.025} = 1.96$
- *Relation to the standardized normal curve*: The normal distribution sets the baseline of 1.96

df	$t_{.025}$
3	3.182
5	2.571
10	2.228
15	2.131
20	2.086
30	2.042
45	2.014
60	2.000
75	1.992
100	1.984
200	1.972
1000	1.962
normal	1.960

4.3b
Construction, Meaning and Interpretation
of the Confidence Interval

Data Sampled from a Stable Process

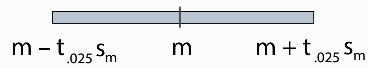
Process mean versus just an arithmetic mean

- Goal is to estimate the value of the population mean, μ
- Can calculate m for *any* set of numbers, such as the weights of 5 Apples and the GMAT scores of 5 MBA candidates
- **Process Mean**: A mean computed from data values that are all sampled from the same process
- For data from different populations, there is an arithmetic mean but different process means
- **Key Concept**: A meaningful sample mean is a process mean
- **Stable Process Model**: *Only* source of variability for all data values of Y is random error at a constant level of variability,
  $$Y_i = \mu + \epsilon_i \quad \textit{where} \quad \sigma_\epsilon = \textit{constant}$$
- If the data values are ordered by time, this *stable process* model (from Section 2.3) should be affirmed before proceeding with the application of the confidence interval

Construction of the Confidence Interval

From sample data to inference . . .

- Begin with three sample statistics: sample size n, sample mean m, and sample standard deviation s
- Calculate the *estimated* standard error: $s_m = \frac{s}{\sqrt{n}}$
- For a 95% confidence level, obtain $t_{.025} \approx 2$ cutoff, $df = n - 1$
- **Margin of Error**, E: Number of estimated standard errors from the sample mean to either end of the confidence interval
- Un-standardize t: Calculate margin of error, $E = (t_{.025})(s_m)$
- Confidence interval: Lower, Upper Bounds: $m - E$ to $m + E$



- Use `lessR` function `ttest()`, here for variable Y: `> ttest(Y)`

Meaning of the Confidence Interval

Classical (frequentist) model of statistics

- Meaning of the analysis follows from what happens over many, many, usually hypothetical, samples
- The value of the population mean, μ, is fixed at a specific value, some constant such as 21 or -4.97
- The estimate of μ, the confidence interval, is a random interval, which randomly varies from sample to sample
  - Sample mean varies from sample to sample, so the position of each interval varies
  - Sample standard deviation varies from sample to sample, so the width of each interval varies
- **Meaning of the 95% confidence interval**: On average, over many samples, each of the same size, 95% of the corresponding confidence intervals contain μ

Probability vs Confidence

The concept of probability applies to random events

- The concept of probability refers to a random outcome of a *future* event
- Stated another way: For a 95% confidence interval, the probability that any one random confidence interval calculated from a future sample *will* contain μ is 95%
- However, *after* the data have been gathered and the one specific confidence interval has been obtained, then either the fixed value μ lies in the obtained interval or it does not
- So, after the confidence interval is computed, the concept of probability does not apply, instead results are correctly stated as: With 95% confidence, . . .
- *Risk cannot be avoided*: In practice only observe a single sample, so do not know if that sample is one of those occasional samples that does *not* contain μ

Simulate Drawing Multiple Samples

The computer provides the data *as if* we have multiple samples

- A sample is just one random sample from a larger population
- **Key Question**: How often does the confidence interval successfully get μ, a hit, or not, a miss, over multiple samples?
- In practice we only have one sample with μ unknown
- Computer simulation allows us to explore directly what would happen with multiple samples

> To obtain many confidence intervals of the mean from repeated sampling of simulated data, all from the same normal population, use the `lessR` function `simCImean`
>
> *Required*: `ns`, number of samples; `n`, size of each sample
>
> *Default*: population mean, `mu=0`
> population standard deviation, `sigma=1`
>
> Optional: `?simCImean`, e.g., `pause=TRUE`

Meaning of the Confidence Level: Illustration

```
> simCImean(ns=50, n=10, mu=65.5, sigma=2.5)
```

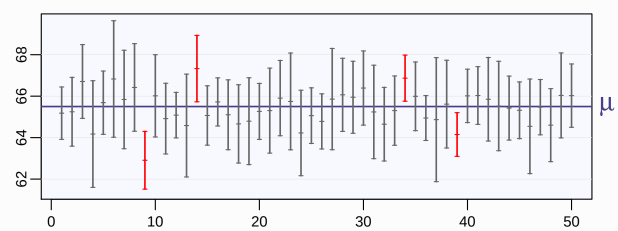


Figure: 50 confidence intervals from 50 samples, $\mu = 65.5$, $\sigma = 2.5$

- For these particular 50 random samples, each of size $n = 10$, 4 of the 95% confidence intervals, or 8%, do not contain μ
- **Key Concept**: Across samples, m, and also s, and so s_m, differ, so the center and width of *each* confidence interval differ

Interpretation of the Confidence Interval of the Mean

Conclusion expressed in jargon free, natural, relaxed English

- Statement of the Confidence Interval: Ship times for 15 shipments from a supplier are recorded and the 95% confidence interval calculated, from 7.21 to 8.97 days
- Interpretation of the Confidence Interval

Interpretation	Purpose
With 95% confidence,	State the Confidence Level
the true average	Unknown population mean, μ
shipping time	Variable of interest
for this supplier	Generalizability of results
is somewhere between	CI is a range, not a single value
$7\frac{1}{5}$ and 9 days	Lower and Upper bounds of CI

4.3c
Application
of the Confidence Interval

Managerial Question

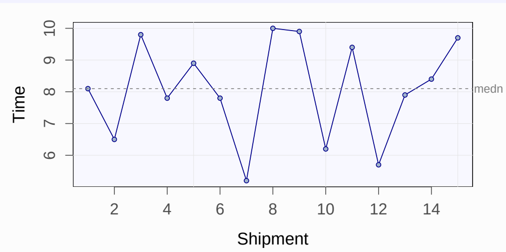
Need for statistical inference

- Criterion: Average supplier ship time no more than 7.5 days
- The expression of this criterion is a specification of the value of the underlying population mean: $\mu \leq 7.5$
- To assess the average delivery time for a supplier is an attempt to estimate the value of μ for that supplier, to generate a forecast about what happens *next year*
- To predict future performance assess past delivery times
- **Problem**: The past, assessed by descriptive statistics, is just one arbitrary sample replete with sampling error
- **Solution**: Use inferential statistics to estimate μ, the basis of the forecast of ship time for a specific supplier
- `> d <- Read("http://lessRstats.com/data/shiptime.csv")`

Evaluate Stability of Process

Process should be stable (in-control)

- Including data from different shipping procedures would not assist in forecasting the future of the *current* process
- The data values are ordered by time, so a line chart is feasible
- `> LineChart(Time, xlab="Shipment")`



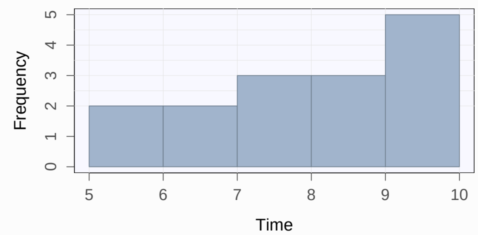
- Time series shows no pronounced pattern for the 15 ship times, so assume stability of μ and σ for all sampled values

Evaluate Normality of Sample Mean via Normality of Data

The histogram of the data

- Sample size here is much less than 30, so evaluate histogram for normality of the data to ensure normality of *m*

```
> Histogram(Time)
```



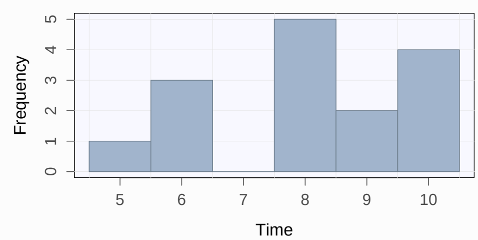
- This histogram is "not particularly" normal
- A better result would be for the last bar, for the 9 to 10 bin, to be centered over the graph, but then this is a small sample

Evaluate Normality of Data, Bin Shift

Consider the bin shift artifact

- Histograms on small data sets are particularly susceptible to histogram artifacts
- Here, just start the same sequence of bins at 4.5 instead of 5

```
> Histogram(Time, bin_start=4.5)
```



- This histogram of the *same* data is "more normal", and so provides support for a normal corresponding sample mean, *m*

Confidence Interval of the Mean: lessR

- lessR function ttest(), or tt(), provides the confidence interval, from the data values directly, or from the summary statistics

```
> ttest(Time)   or   > ttest(n=15, m=8.09, s=1.59)
```

```
--- Description ---

n = 15,   mean = 8.09,   sd = 1.59


--- Inference ---

t-cutoff: tcut =  2.145
Standard Error of Mean: SE =  0.41

Margin of Error for 95% Confidence Level:  0.88
95% Confidence Interval for Mean:  7.21 to 8.97
```

Confidence Interval of the Mean: Traditional Notation

Write the equations manually

- **Data Summary**. $n = 15$, $m = 8.087$ days, $s = 1.587$ days
- **Estimated standard error of the mean**.
  $s_m = \frac{s}{\sqrt{n}} = \frac{1.587}{\sqrt{15}} = 0.410$ days
- **$t_{.025}$-cutoff**. $df = 15 - 1 = 14$, $\quad t_{.025} = 2.145$
- **Margin of Error**. $E = (t_{.025})(s_m) = 2.145(0.410) = 0.879$
- **Confidence interval**.
  Lower Bound: $m - E = 8.087 - 0.879 = 7.21$ days
  Upper Bound: $m + E = 8.087 + 0.879 = 8.97$ days

Interpret the Confidence Interval

Verbal Restatement

- 95% confidence interval for average ship time is 7.21 to 8.97 days
- This statement is *not* an interpretation because it presumes the intended audience understands the meaning of the technical phrase "confidence interval"
- Instead this statement just restates the confidence interval with words

Meaningful Interpretation

- With 95% confidence, the true average shipping time for this supplier is somewhere between $7\frac{1}{5}$ and 9 days

Conclusion

Analysis of Descriptive Statistics

- For last year's 15 deliveries, average delivery time was 8.09 days, yet management will only accept an average of 7.5 days

Analysis of Inferential Statistics

- An average delivery time as low as 7.21 days is plausible

Managerial Decision

- Before finding another supplier, provide more opportunity to demonstrate a sufficiently small delivery time, either with more data, or maybe a qualitative analysis and potential improvement of the supplier's internal process for shipping

- The End

Index

Subtract 2 from each listed value to get the Slide #