Confidence Intervals

GrowingKnowing.com © 2011

Estimates

- We are often asked to predict the future!
  - When will you complete your team project?
  - When will you make your first million dollars?
  - When will you clean the dishes?

- You can give a **point estimate** or **interval estimate**
  - Point estimate: Project will be done Friday at 2
  - Interval estimate: Project will be done this week (7 day range)
- The smart answer is the interval estimate
  - Because it includes a range to allow for variability
    - Life has surprises, illness, accidents…. standard deviations
  - For a point estimate, 2.01 pm, if you're late, … trouble!

Confidence Level

- Whatever promise you make
  - to your boss, team, … interval or point estimate
- If they are smart, their next question will be

- **HOW CONFIDENT ARE YOU?**

- If you say 99% confident level, they won't worry
- If you say 20% confident , …
  - Trust declines, not so good, …. 'we have to talk?'

Confidence Interval vs. Confidence Level

- Confidence **Interval**
  - **Range**: I'll clean the dishes between Monday at 2pm and next September.
    - Confidence interval is a *range* of values that is expected to include an unknown population parameter based on your sample.
    - Note: You have an **upper and lower** answer
- Confidence **Level**
  - **Level** : I 'm 99% confident dishes will be clean in this interval.
    - Confidence level is *how likely the value* will *fall within your* confidence *interval*.
- Confidence level and interval are different.
  - Interval is a range. Confidence level is a percentage.

Some important words

- **Level of significance**
- **Alpha**
- **Confidence Level**

- These words are all talking about the *same* idea.
- Alpha is level of significance.
  - Alpha just a shorter word for level of significance.
  - Alpha is the complement of Confidence Level.

- **Alpha = 1 – confidence level**
- **Confidence = 1 - alpha**
  - What is alpha if Confidence level is 99%,
    - Alpha = 1 - .99 = .01
  - What is alpha if Confidence level is 95%,
    - Alpha = 1 - .95 = .05
  - What is the confidence level if Alpha is .10,
    - then Confidence Level is 1 - .1 = 90%
  - Alpha is 5%, what is the level of significance?
    - .05 or 5%. Alpha is the level of significance.
  - Level of significance is 5%, what is confidence level?
    - Confidence level = 1 – level of significance = 95%

Important calculations

- Standard Error for Means
  = Standard deviation divided by square root of sample size (n)

  Standard Error = $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

- Margin of Error (E)
  = z multiplied by standard error (σ_x)

  Margin of Error = $E = z\sigma_{\bar{x}}$

- Confidence Interval
  - **Upper interval** = mean + margin of error = $\bar{x} + z\sigma_{\bar{x}}$
  - **Lower interval** = mean – margin of error = $\bar{x} - z\sigma_{\bar{x}}$

Test

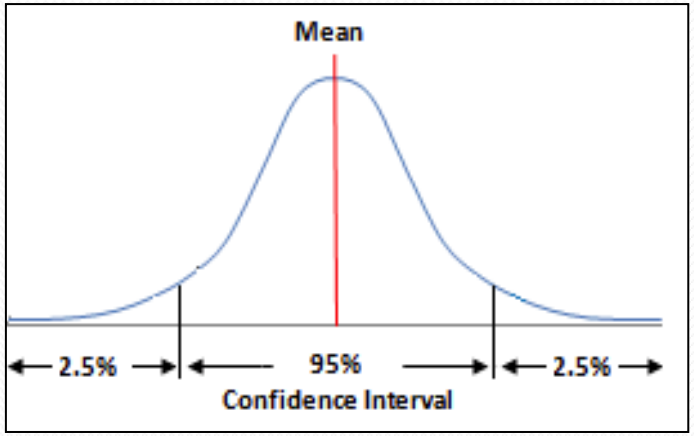
- Standard deviation (S.D.) is 10, sample size (n) is 100. What is the standard error (Std error) ?
  - Std error = $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ = S.D. / Square root (n)
  - Std error = $10/\sqrt{100} = 10/10 = 1$
- What is the standard error if the mean is 45, standard deviation is 4, and sample size is 25?
  - Std error = S.D. / $\sqrt{n}$
  - $= 4/\sqrt{25} = 4/5 = .8$

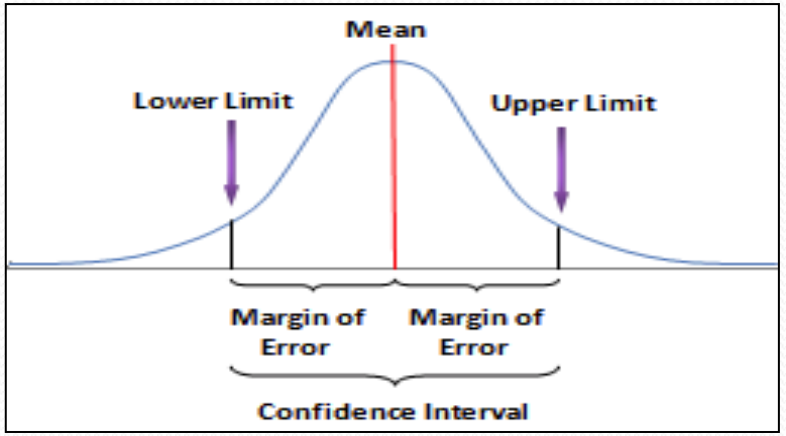
Confidence

- Confidence is a 2 tail calculation (upper and lower values)
- You can make an error by predicting too low or too high.
  - 2 ways to make a mistake.
- Example: you predict the Maple Leafs will lose 150 to 200 games
  - You could be too low, they lose the next 350 games.
  - You could be too high, they lose only the next 144 games.
  - Either too high or too low indicate your interval was not correct.
- A realistic example, you want to sell your stock at just the right price. Your confidence level says sell between 101 and 109 dollars.
  - If you are wrong and it goes higher, you lose profits.
  - If are wrong and it never reaches 101, you lose opportunity.
  - Two tail calculation.

Error is too high or too low

- For 95% confidence, you expect 95% of your samples will fall between lower and upper interval values.
- Alpha is 5%, so 5% the samples will be outside the interval.
- 5% divide by 2, so 2.5% of samples will be too high, and 2.5% will be too low.





Calculating z

- Example: confidence level is 95%
  - Probability is (confidence level + alpha / 2)
  - Probability is (.95 + .05/2) = .9750
    - z value for probability .9750 = 1.96
  - z = 1.96
- Link to Normal table
  - www.growingknowing.com/GKStatsBookNormalTable2.html

Calculating z

- Example: confidence level is 98%
  - Probability (confidence level + alpha /2)
    - Probability = .98 + .02/2 = .99
  - z value for probability .9900 = 2.33

What is the right amount of confidence?

- Most studies use 90%, 95%, and 99% confidence levels
  - How do you decide which level to pick?
- For $1 million dollars, this is an important decision, so be sure to set a 99% confidence level.
- For a $1 bet, not so important, 90% confidence is fine.

- To speed up your tests and **reduce errors**, memorize the popular confidence levels.
  - 90% confidence z = 1.64
  - 95% confidence z = 1.96
  - 99% confidence z = 2.58

Calculating Margin of Error (E)

- What is margin of error (E) if n = 75, std. deviation (S.D.) is 45.08, and confidence level is 95%?
- Step 1: Calculate standard error
  - Std error = S.D. / Square root (n)
  - $= 45.08 / \sqrt{75}$
  - Std error = 45.08/ 8.660254 = 5.205
- Step 2: Calculate z
  - Z = 1.96
  - we were just told to memorized 90%, 95%, and 99%.
- Step 3: Margin of Error (E) = z(std error)
  - E = 1.96(5.205) = 10.2

Confidence intervals

- Step 4: Calculate confidence intervals

- If the margin of error (E) is 10.2 and mean is 92, what are the confidence intervals?
  - Upper interval = mean + E
  - Lower interval = mean – E
- Upper is 92 + 10.2 = 102.2
- Lower is 92 – 10.2 = 81.8

Example

- What is the Confidence Interval where n = 75, mean = 61, S.D. = 8.54, and confidence level =96%?

- Step 1: Std Error = S.D./ square root (n)
- Step 2: Find z
- Step 3: Margin of error E = z(std error)
- Step 4: Confidence intervals Mean +/- E

- **What is the Confidence Interval where sample is 75, mean = 61, std deviation (S.D.) = 8.54, and confidence level of 96%?**

- Step 1: Std Error = S.D./ square root (n)
  - $= 8.54 / \sqrt{75} = 8.54/8.66 = .986$
- Step 2: Find z
  - Probability (confidence level + alpha /2)
  - $(.96 + .04/2) = .9800$
  - z = 2.055 (half way between .9798 and .9803)
- Step 3: Margin of error E = z(std error)
  - $= 2.055(.986) = 2.026$
- Step 4: Confidence intervals Mean +/- E
  - Upper interval = mean + 2.026 = 61 + 2.026 = 63.03
  - Lower interval = mean - 2.026 = 61 – 2.026 = 58.97

Practice

- Go to website and practice Confidence Interval
  - Difficulty level 1 only
  - To complete level 2, you need the proportion lesson

Small Samples Confidence

- If your sample size (n) is less than 30, you cannot use the normal distribution table (see Central Limit Theory)
- For small samples, use the Student t table
  - t table is more robust, works well with data that is not perfectly normal.
- The t table uses alpha and degrees of freedom (df).

$$\textbf{degrees of freedom (df)} = n - 1$$

- The only difference in calculating small sample confidence and large samples is the t table, all other calculations and steps are the same.

Degrees of freedom

- Put your hand on your head, now take a sample of 21, and a mean of 22, and calculate the degrees of freedom?
  - $df = n - 1$
  - $df = 21 - 1 = 20$

- Taxes, work, bosses, parents, …. is freedom an illusion?

Using the t table

- Calculate t for 90% confidence level using sample 16?
  - Degrees of freedom = n – 1
    - df = 16 – 1 = 15
  - Alpha = 1 – confidence level
  - alpha = 1 - .9 = .10
- Lookup 2 tail t table for .10 and df = 15, t = 1.753
  - http://www.growingknowing.com/GKStatsBook.php?topic=StudentTTable

Example

- What is the Confidence Interval where **n = 16**, mean = 53, S.D. = 14.84, and confidence level =90%?

- Step 1: Std Error = S.D./ square root (n)
- Step 2: **Calculate t**
- Step 3: Margin of error E = t(std error)
- Step 4: Confidence intervals Mean +/- E

- **What is the Confidence Interval where n = 16, mean = 53, S.D. = 14.84, and confidence level =90%?**

- Step 1: Std Error = S.D./ square root (n)
  - = 14.84 / $\sqrt{16}$ = 14.84/4 = 3.71
- Step 2: Calculate t
  - df = n – 1 = 16 – 1 = 15
  - alpha = 1 - .9 = .1
  - Lookup 2 tail t table .10 with df 15 = 1.753
- Step 3: Margin of error E = t(std error)
  - = 1.753(3.71) = 6.50
- Step 4: Confidence intervals Mean +/- E
  - Upper interval = mean + 6.50 = 53 + 6.50 = 59.5
  - Lower interval = mean - 6.50 = 53– 6.50 = 46.5

- **You survey vampires to see if they prefer eating Italian? What is the Confidence Interval if n = 17, mean = 60, S.D. = 7.8, and confidence level =95%?**

- Step 1: Std Error = S.D./ square root (n)
  - $= 7.8 / \sqrt{17} = 7.8/4.1231 = 1.89$
- Step 2: Calculate t
  - df = n – 1 = 17 – 1 = 16
  - alpha = 1 - .95 = .05
  - Lookup 2 tail t table .05 with df 16 = 2.120
- Step 3: Margin of error E = t(std error)
  - = 1.89(2.12) = 4.01
- Step 4: Confidence intervals Mean +/- E
  - Upper interval = mean + 4.01 = 60+ 4.01 = 64.01
  - Lower interval = mean - 4.01 = 60– 4.01= 55.99

Confidence levels

- People like a high confidence level but small interval.
  - I'm 95% confident I'll deliver between Monday and August. Will you be home?
- It is easy to have small interval with small confidence level
  - I'll deliver Friday 12.00 to 12:05 am, I'm 50% confident.
- If you increase the confidence level, the interval gets larger
- How can you get both high confidence and small intervals?
- Look at the formula, best way is increase the sample size.

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

Practise

- Go to the website, do Small Sample Confidence.