

# Statistik Bisnis

Week 9

Confidence Interval Estimation

# Agenda

Time	Activity
20 minutes	Point and Interval Estimate
40 minutes	Confidence Interval Estimate for the Mean ( $\sigma$ Known)
40 minutes	Confidence Interval Estimate for the Mean ( $\sigma$ Unknown)
40 minutes	Confidence Interval Estimate for the Proportion
30 minutes	Determining Sample Size
30 minutes	Exercise

# Outline

- Point and Interval Estimate
- Confidence Interval Estimate for the Mean ( $\sigma$  Known)
- Confidence Interval Estimate for the Mean ( $\sigma$  Unknown)
- Confidence Interval Estimate for the Proportion
- Determining Sample Size

# What for?

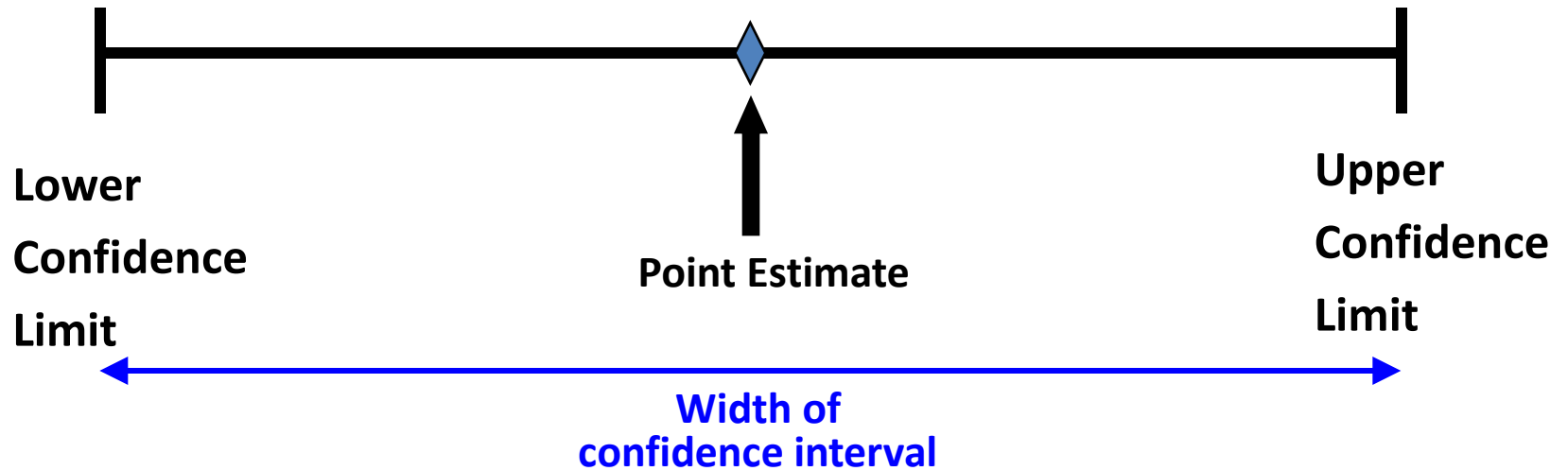


# Estimation

- A **point estimate** is the value of a single sample statistic, such as a sample mean.
- A **confidence interval estimate** is a range of numbers, called an interval, constructed **AROUND THE POINT ESTIMATE.**

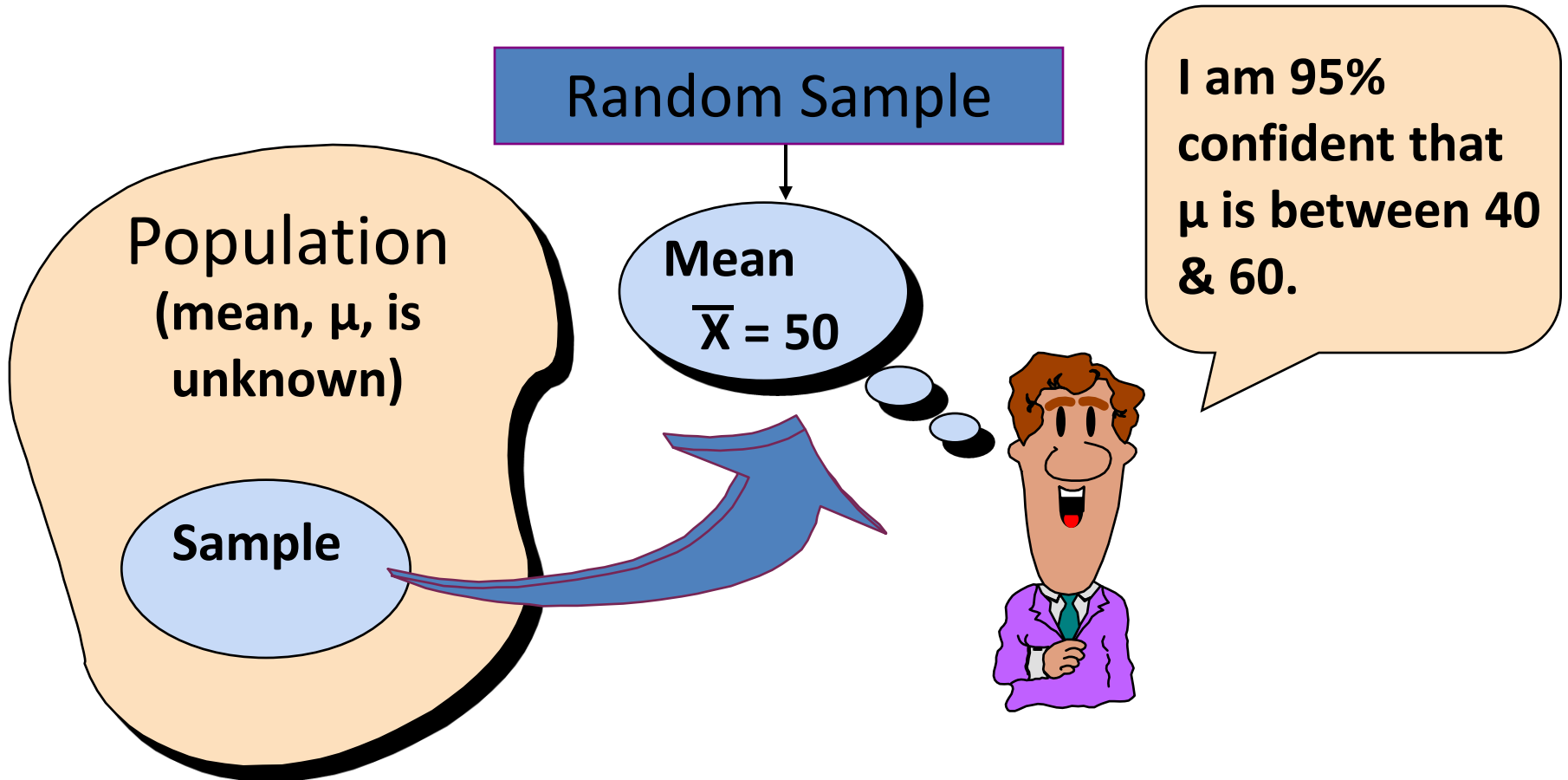


# Estimation



We can estimate a Population Parameter ...		with a Sample Statistic (a Point Estimate)
Mean	$\mu$	$\bar{X}$
Proportion	$\pi$	$p$

# Estimation Process



# General Formula

- The general formula for all confidence intervals is:

$$\text{Point Estimate} \pm (\text{Critical Value})(\text{Standard Error})$$

Where:

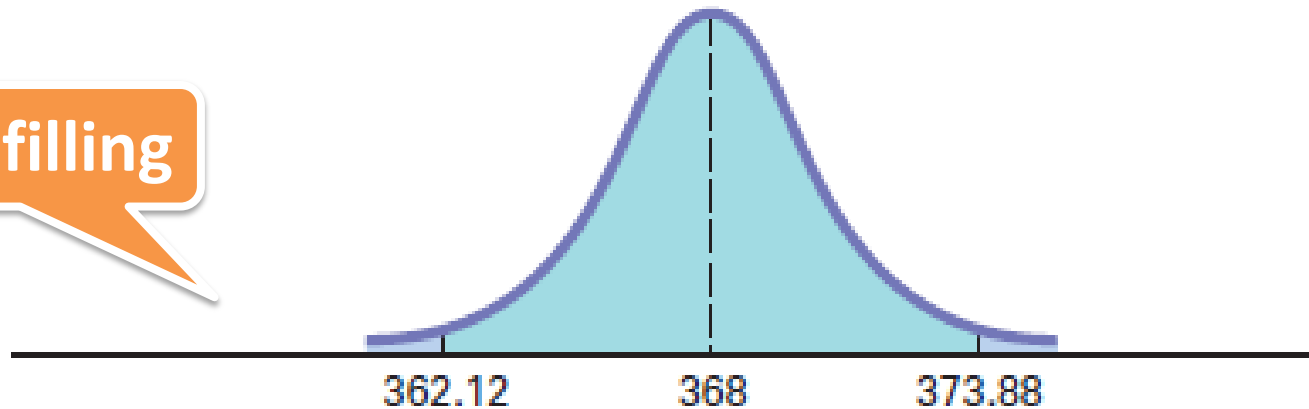
- **Point Estimate** is the sample statistic estimating the population parameter of interest
- **Critical Value** is a table value based on the sampling distribution of the point estimate and the desired confidence level
- **Standard Error** is the standard deviation of the point estimate



**CONFIDENCE INTERVAL ESTIMATE  
FOR THE MEAN ( $\sigma$  KNOWN)**

# Population Mean vs. Sample Mean

cereal-filling



$\bar{X}_1 = 362.3$



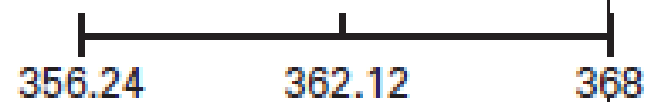
$\bar{X}_2 = 369.5$



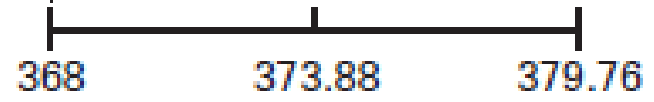
$\bar{X}_3 = 360$



$\bar{X}_4 = 362.12$



$\bar{X}_5 = 373.88$



# Level of Confidence

$$(1 - \alpha) \times 100\%$$

Where:

$\alpha$  is the proportion in the tails of the distribution that is outside the confidence interval.

# Assumptions

- Population standard deviation ( $\sigma$ ) is known
- Population is normally distributed
- If population is not normal, use large sample

# CONFIDENCE INTERVAL FOR THE MEAN ( $\sigma$ KNOWN)

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

or

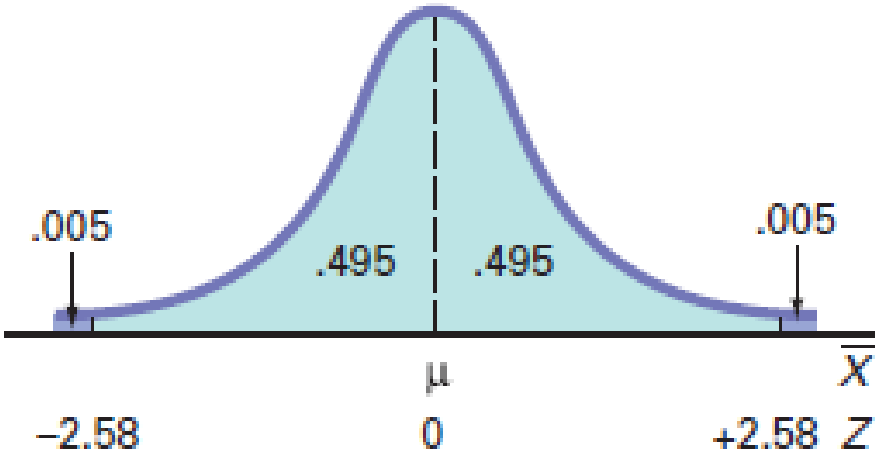
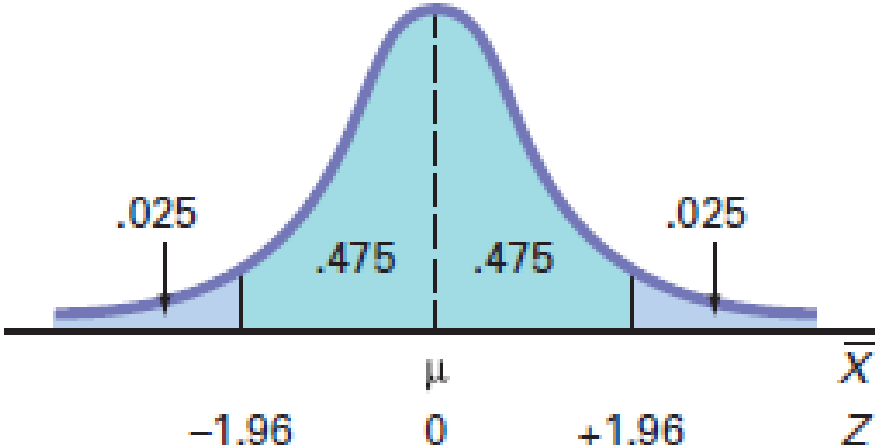
$$\bar{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

# Critical Value

- The value of  $Z_{\alpha/2}$  needed for constructing a confidence interval is called the **critical value** for the distribution.

LOC	$\alpha$	$\alpha/2$	$Z_{\alpha/2}$
99%	0.01	0.005	2.58
95%	0.05	0.025	1.96
90%	0.1	0.05	1.645

# Increasing Level of Confidence



# Increasing Level of Confidence

- Any increase in the level of confidence is achieved only by widening (and making less precise) the confidence interval.
- There is no “free lunch” here. You would have more confidence that the population mean is within a broader range of values; however, this might make the interpretation of the confidence interval less useful.



# Do You Ever Truly Know $\sigma$ ?

- Probably not!
- In virtually all real world business situations,  $\sigma$  is not known.
- If there is a situation where  $\sigma$  is known then  $\mu$  is also known (since to calculate  $\sigma$  you need to know  $\mu$ .)
- If you truly know  $\mu$  there would be no need to gather a sample to estimate it.

**EXAMPLE**

# Paper Manufacture

Sebuah perusahaan kertas memiliki mesin yang beroperasi terus menerus sepanjang shift. Kertas yang diproduksi diharapkan memiliki panjang rata-rata 11 inci dan simpangan baku 0,02 inci. pada waktu-waktu tertentu dalam shift tersebut, diambil sebuah sampel untuk menentukan apakah panjang rata-rata kertas masih 11 inci atau telah terjadi sesuatu yang salah pada mesin sehingga perlu dilakukan penyesuaian pada mesin tersebut. Anda memilih sampel acak 100 lembar kertas, dan panjang rata-rata sampel kertas tersebut adalah 10.998 inci.

# Paper Manufacture

- a. Buatlah pendugaan interval dengan tingkat kepercayaan 95% untuk panjang rata-rata populasi kertas yang diproduksi pada shift tersebut.
- b. Buatlah pendugaan interval dengan tingkat kepercayaan 99% untuk panjang rata-rata populasi kertas yang diproduksi pada shift tersebut.

**CONFIDENCE INTERVAL ESTIMATE  
FOR THE MEAN ( $\sigma$  UNKNOWN)**

# Confidence Interval For The Mean ( $\sigma$ Unknown)

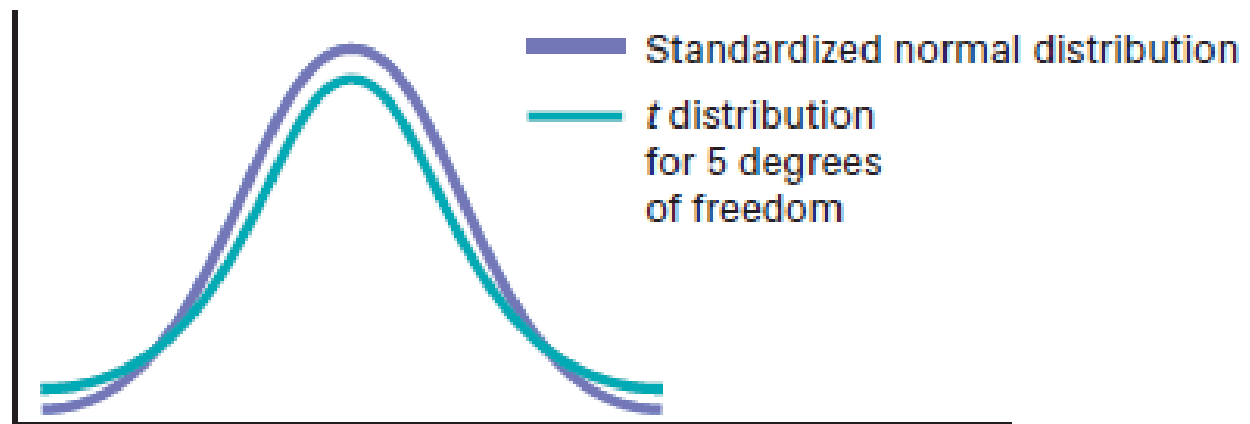
$$\bar{X} \pm t_{\alpha/2, df} \frac{S}{\sqrt{n}}$$

or

$$\bar{X} - t_{\alpha/2, df} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2, df} \frac{S}{\sqrt{n}}$$

# Student's t Distribution

- William S. Gosset
- Assumes that the random variable  $X$  is *normally distributed*
- The t Table



# Student's t Distribution

- The t is a family of distributions
- The  $t_{\alpha/2}$  value depends on **degrees of freedom (d.f.)**
  - Number of observations that are free to vary after sample mean has been calculated

$$\text{d.f.} = n - 1$$



# Degrees of Freedom (df)

**Idea:** Number of observations that are free to vary after sample mean has been calculated

**Example:** Suppose the mean of 3 numbers is 8.0

Let  $X_1 = 7$   
Let  $X_2 = 8$   
What is  $X_3$ ?



If the mean of these three values is 8.0, then  $X_3$  **must be 9** (i.e.,  $X_3$  is not free to vary)

Here,  $n = 3$ , so degrees of freedom =  $n - 1 = 3 - 1 = 2$

(2 values can be any numbers, but the third is not free to vary for a given mean)

**EXAMPLE**

# Electric Insulators

Sebuah perusahaan manufaktur memproduksi isolator listrik. Jika isolator pecah saat digunakan, bisa terjadi hubung singkat. Untuk menguji kekuatan isolator, dilakukan tes destruktif untuk mengetahui tekanan yang dibutuhkan hingga isolator tersebut pecah. Tekanan yang dibebankan pada isolator, tepat sebelum isolator tersebut pecah kemudian dicatat. Percobaan ini dilakukan pada 30 isolator yang dipilih secara acak.

# Electric Insulators

Berikut data yang telah dikumpulkan melalui percobaan tersebut:

1,8704	1,728	1,656	1,610	1,634	1,784	1,522	1,696	1,592	1,662
1,866	1,764	1,734	1,662	1,734	1,774	1,550	1,756	1,762	1,866
1,820	1,744	1,788	1,688	1,810	1,752	1,680	1,810	1,652	1,736

Untuk menganalisis data, anda perlu membentuk pendugaan interval dengan tingkat kepercayaan 95% untuk mengetahui tekanan rata-rata yang dibutuhkan untuk memecahkan isolator

# **INTERVAL ESTIMATE FOR THE PROPORTION**

# Confidence Interval For Proportion

$$p \pm Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

or

$$p - Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p + Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

**Note: must have  $np > 5$  and  $n(1-p) > 5$**

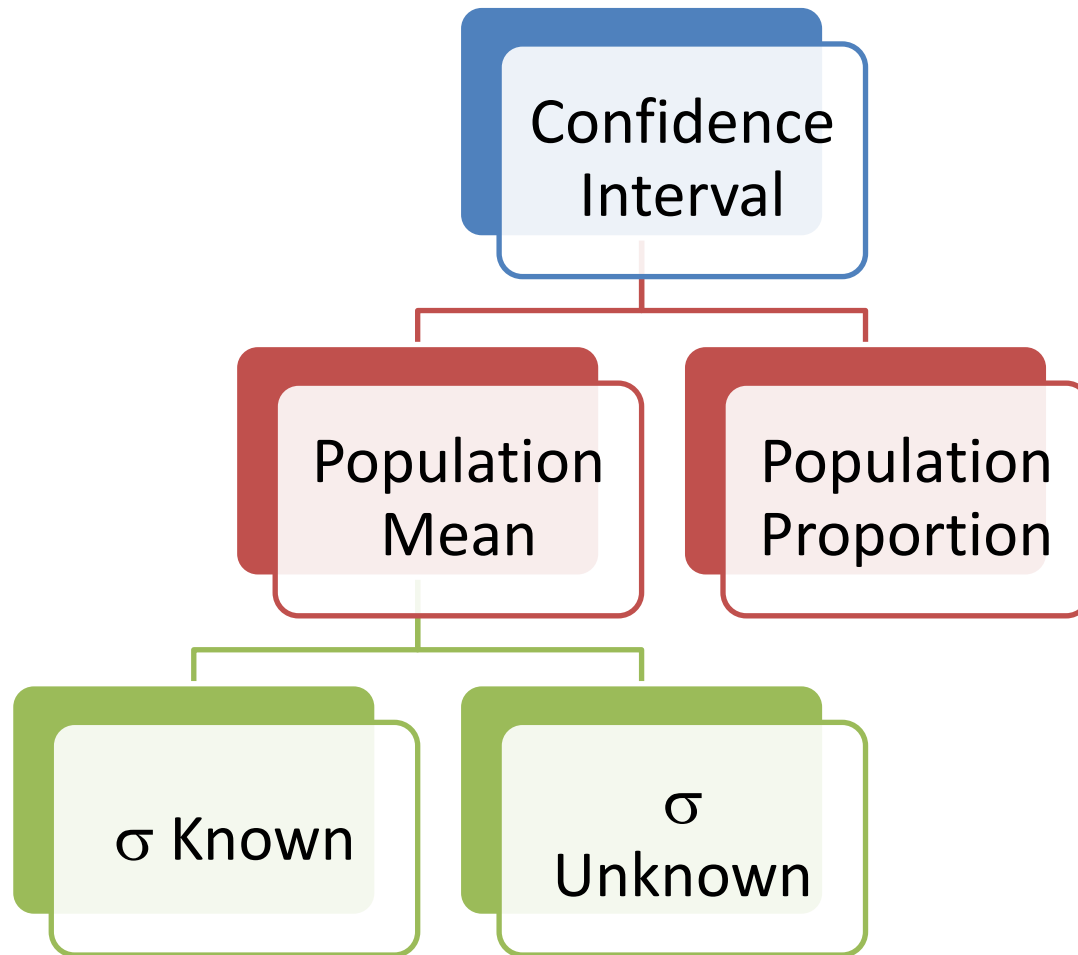
**EXAMPLE**

# Newspaper

Manajer operasi sebuah perusahaan koran terkemuka ingin mengestimasi proporsi koran yang memiliki cacat cetak. Anda mengumpulkan data dengan memilih 200 sampel acak dari semua koran yang dicetak pada satu hari tertentu, melakukan pengecekan dan mengidentifikasi 35 koran yang memiliki cacat cetak. Untuk menganalisis data, anda perlu membangun estimasi interval dengan tingkat kepercayaan 90% dari proporsi koran yang memiliki cacat cetak pada hari tersebut. Interpretasikan hasil pendugaan interval tersebut!

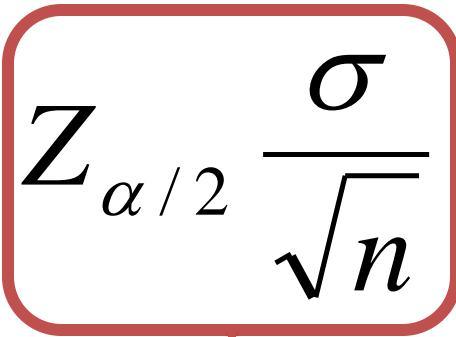


# Summary on Confidence Interval



# **DETERMINING SAMPLE SIZE**

# Sample Size For The Mean

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$




Sampling Error (e)

# Sample Size For The Mean

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$n = \frac{Z_{\alpha/2}^2 \sigma^2}{e^2}$$

# Sample Size For The Population

$$e = Z_{\alpha/2} \sqrt{\frac{\pi(1-\pi)}{n}}$$

$$n = \frac{Z_{\alpha/2}^2 \pi(1-\pi)}{e^2}$$

**EXAMPLE**

# Electric Insulators

Kembali pada soal latihan isolator listrik sebelumnya, misalkan anda ingin mengestimasi dengan tingkat kepercayaan 95%, tekanan rata-rata yang dibutuhkan untuk memecahkan isolator, dengan tingkat kesalahan  $\pm 25$  pounds. Berdasarkan penelitian yang dilakukan pada tahun sebelumnya, anda percaya bahwa simpangan baku populasinya adalah 100 pounds. Tentukan ukuran sampel yang dibutuhkan.

# Office Worker

Anda ingin memiliki tingkat kepercayaan 90% dalam mengestimasi proporsi pekerja kantoran yang membalas email dalam 1 jam dengan tingkat kesalahan  $\pm 0.05$ . Karena anda tidak pernah melakukan penelitian sejenis sebelumnya, tidak ada informasi masa lalu yang dapat digunakan. Tentukan ukuran sampel yang diperlukan.



**EXERCISE**

## 8.9 (cont'd)

Manager sebuah toko cat ingin mengestimasi volume aktual pada 1 galon kaleng cat yang ia beli dari produsen cat ternama. Produsen tersebut menyatakan bahwa deviasi standar volume cat tersebut adalah 0,02 galon. Sebuah sampel acak berukuran 50 kaleng dipilih, dan rata-rata volume cat dalam kaleng berukuran 1 galon tersebut adalah 0,995 galon.

## 8.9

- a. Buatlah pendugaan interval dengan tingkat kepercayaan 99% untuk rata-rata populasi volume cat pada kaleng berukuran 1 galon.
- b. Berdasarkan hasil tersebut, apakah manajer tersebut dapat mengajukan keluhan kepada produsen cat? Kenapa?
- c. Buatlah pendugaan interval dengan tingkat kepercayaan 95%. Bagaimana hal ini mempengaruhi jawaban anda pada poin (b)?

## 8.15

Sebuah toko alat tulis ingin mengestimasi nilai retail rata-rata dari kartu ucapan yang dimilikinya. Sampel berukuran 100 kartu ucapan mengindikasikan nilai rata-rata \$2,55 dan deviasi standar \$0,44.

- a. Dengan mengasumsikan distribusi normal, buatlah pendugaan interval dengan tingkat kepercayaan 95% untuk nilai rata-rata kartu ucapan di gudang toko tersebut.
- b. Misalkan terdapat 2.500 kartu ucapan di gudang toko tersebut. Bagaimana hasil pada poin (a) dapat membantu pemilik toko untuk mengestimasi total nilai kartu ucapan dalam gudang?

## 8.28 (cont'd)

Sebuah penyedia layanan TV kabel ingin mengestimasi proporsi rumah tangga yang mau memasang tambahan line jika mereka memberi pengurangan biaya pemasangan. Data dikumpulkan dari sampel acak 500 rumah tangga. Hasilnya menunjukkan bahwa 135 rumah tangga mau memasang tambahan line jika ada pengurangan biaya pemasangan.

## 8.28

- a. Buatlah pendugaan interval dengan tingkat kepercayaan 99% untuk proporsi populasi rumah tangga yang mau melakukan penambahan line TV kabel.
- b. Bagaimana manajer yang bertanggung jawab terhadap program promosi dapat menggunakan hasil pada poin (a)?

## 8.38 (cont'd)

Sebuah survei direncanakan untuk mengetahui rata-rata pengeluaran tahunan keluarga karyawan di sebuah perusahaan untuk biaya kesehatan. Manajemen perusahaan ingin memiliki kepercayaan 95% bahwa sampel yang terpilih akan berada pada  $\pm \$50$  dari rata-rata pengeluaran tahunan tersebut. Penelitian terdahulu menunjukkan bahwa deviasi standarnya adalah \$400.

## 8.38

- a. Berapa besar sampel yang dibutuhkan?
- b. Jika manajemen ingin benar dalam rentang  $\pm \$25$ , berapa jumlah karyawan yang perlu dipilih?



## 8.16 (cont'd)

Southside Hospital di Bay Shore, New York, biasa melakukan stress test untuk mempelajari otot jantung setelah seseorang mendapatkan serangan jantung. Rumah sakit ini ingin mengurangi waktu pelayanan, yaitu waktu yang dibutuhkan dari tes dilakukan hingga hasil tes keluar. Biasanya rata-rata waktu pelayanan tersebut adalah 68 jam. Setelah melakukan perubahan pada proses pelaksanaan tes, dipilih sampel 50 stress test. Dalam sampel ini didapat rata-rata waktu pelayanan 32 jam, dengan simpangan baku 9 jam.

## 8.16

- a. Buatlah pendugaan interval dengan tingkat kepercayaan 95% untuk rata-rata populasi waktu pelayanan.
- b. Interpretasikan interval pada poin (a).
- c. Apakah menurut anda proyek peningkatan kualitas pelayanan tersebut sukses?

## 8.47 (cont'd)

Pada sebuah studi pada 500 orang eksekutif, 315 menyatakan bahwa perusahaan secara informal memonitor sosial media untuk tetap teratas sehubungan dengan informasi yang berhubungan dengan perusahaan mereka.

## 8.47

- a. Buatlah pendugaan interval dengan tingkat kepercayaan 95% untuk proporsi perusahaan yang memonitor media sosial secara informal tersebut.
- b. Interpretasikan interval yang didapat pada poin (a).
- c. Jika anda ingin melakukan penelitian lanjutan untuk mengestimasi proporsi perusahaan yang memonitor media sosial secara informal dalam rentang  $\pm 0,01$  dengan tingkat kepercayaan 95%, berapa banyak eksekutif yang perlu anda survei?

## 8.46 (cont'd)

Sebuah survei dilakukan pada 1.000 orang pada bulan Maret 2009 mengenai “*green practices*.” Respon untuk pertanyaan apakah hal yang paling bermanfaat yang dapat dilakukan untuk lingkungan: 28% menyatakan dengan membeli energi terbarukan, 19% menyatakan dengan menggunakan moda transportasi ramah lingkungan, dan 7% menyatakan dengan mengurangi atau meminimalisir pengemasan.

## 8.46

Buatlah pendugaan interval dengan tingkat kepercayaan 95% bagi proporsi populasi yang menganggap hal yang paling bermanfaat yang dapat dilakukan untuk lingkungan adalah

- a. membeli energi terbarukan.
- b. menggunakan moda transportasi ramah lingkungan.
- c. mengurangi atau meminimalisir pengemasan.
- d. Anda diminta untuk meng-update hasil penelitian ini. Tentukan ukuran sampel yang dibutuhkan, dengan tingkat kepercayaan 95%, proporsi populasi pada poin (a) hingga (c) dalam  $\pm 0,02$

See you next week...

**THANK YOU**