twitter: @kh_notodiputro

IPB University
Bogor Indonesia

Department of Statistics
Study Program in Statistics and Data Science

# Pengujian Hipotesis


**Prof. Dr. Ir. Khairil Anwar Notodiputro, MS**

*email: khairil@apps.ipb.ac.id*

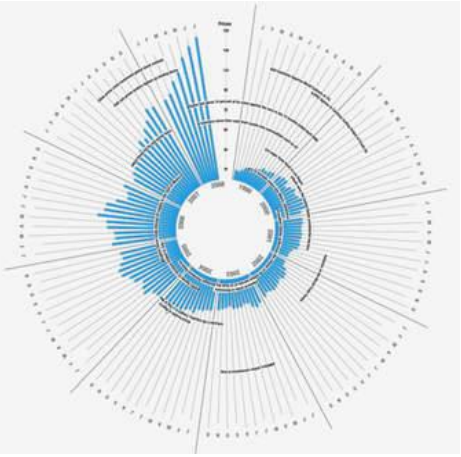
**Ketua Program Studi**
**Statistika dan Sains Data**



# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



## Motivasi

- Tukang ramal percaya bhw posisi planet dan bulan pada saat Anda lahir menentukan kepribadian Anda.
- Apakah ada bukti ilmiah tentang benar/salahnya tukang ramal itu?
- Kalu ingin diuji kemampuan peramal tersebut, bagaimana caranya?

**Example 1**

### Are Astrology Predictions Better than Guessing?

**Picture the Scenario**

Astrologers believe that the positions of the planets and the moon at the moment of your birth determine your personality traits. But have you ever seen any scientific evidence that astrology works? One scientific test of astrology used the following experiment: Each of 116 volunteers was asked to give his or her date and time of birth. From this information, an astrologer prepared each subject's horoscope based on the positions of the planets and the moon at the moment of birth. Each volunteer also filled out a California Personality Index survey. Then the birth data and horoscope for one subject, together with the results of the personality survey for that individual and for two other participants randomly selected from the experimental group, were given to an astrologer. The astrologer was asked to predict which of the three personality charts matched the birth data and horoscope for the subject.[1]

Let $p$ denote the probability of a correct prediction by an astrologer. Suppose an astrologer actually has no special predictive powers, as would be expected by those who view astrology as "quack science." The predictions then merely correspond to random guessing, that is, picking one of the three personality charts at random, so $p = 1/3$. However, the participating astrologers claimed that $p > 1/3$. They felt they could predict better than with random guessing.

**Questions to Explore**

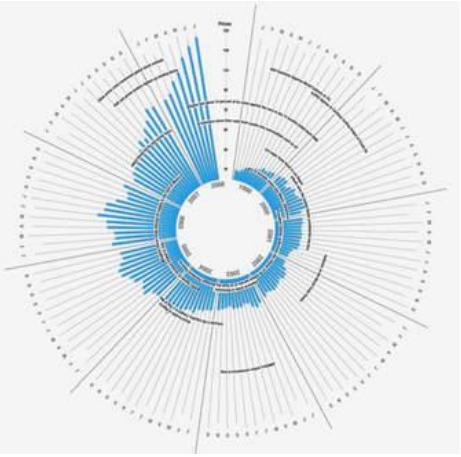
- How can we use data from such an experiment to summarize the evidence about the claim by the astrologers?
- How can we decide, based on the data, whether the claim is plausible?

[1] S. Carlson, *Nature*, vol. 318, pp. 419–425, 1985.

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



Kemampuan peramal bisa diuji dengan cara sbb:

116 sukarelawan/wati menyerahkan tgl, bln, dan th kepada tukang ramal. Pada saat yg sama 116 orang itu mengisi data tentang sikap, pikiran dan perilaku yang dapat dianalisis untuk mengetahui kepribadiannya oleh psikolog.

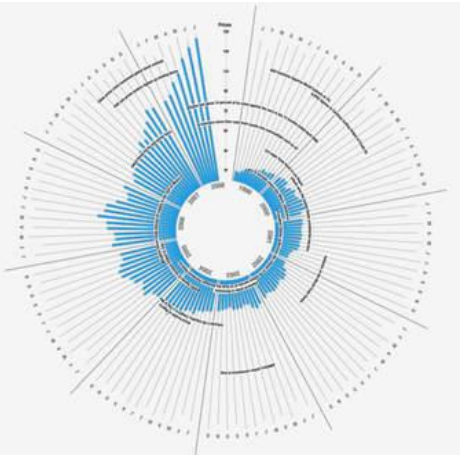

- Tgl bln dan thn lahir sukarelawan diberikan kpd tukang ramal. Lalu peramal meramal sikap, pikiran dan perilaku sukarelawan.
- Seberapa tepat peramal dapat menentukan kepribadian seseorang sesuai hasil analisis kepribadian?
- Tukang ramal tepat jika hasilnya sesuai dengan isian sukarelawan.
- Misal $p$ = peluang benar. Jika $p = 1/3$, artinya menebak (acak).
- Jika $p > 1/3$, artinya punya kemampuan meramal.

**Kemampuan meramal tsb dpt dievaluasi melalui PENGUJIAN HIPOTESIS…!**

# Pengujian Hipotesis

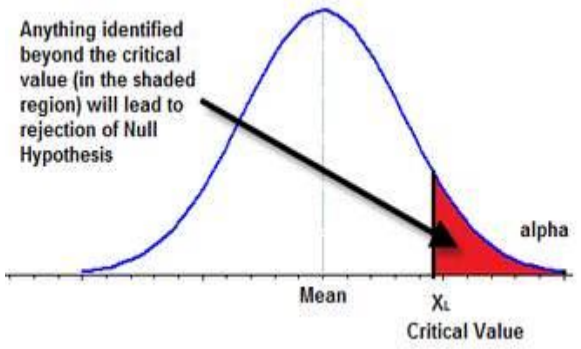
Hypothesis Testing

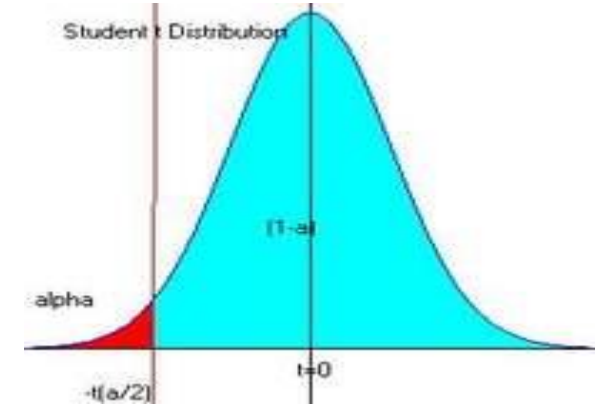
IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



## Apa itu pengujian hipotesis?

- Dalam statistika, hipotesis adalah pernyataan tentang populasi.
- Biasanya merupakan klaim thd nilai parameter populasi.
- Pengujian hipotesis merupakan metode untuk mengambil kesimpulan tentang parameter populasi.
- Ingat tentang peramal tadi, p > 1/3 berarti peramal punya kemampuan:
  - Jika peramal menebak benar 10 dari 116 data orang maka kita ragukan peramal itu → hanya nebak
  - Jika peramal menebak benar 100 dari 116 data orang itu maka kita yakin dengan kemampuan peramal tsb → mungkin tdk menebak

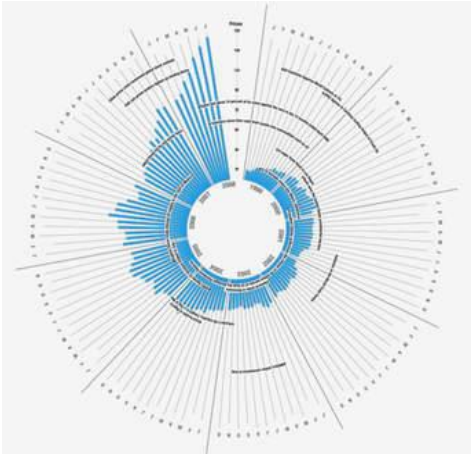


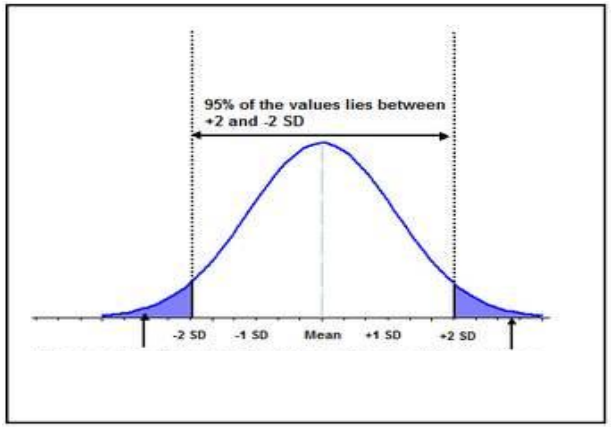

# Pengujian Hipotesis

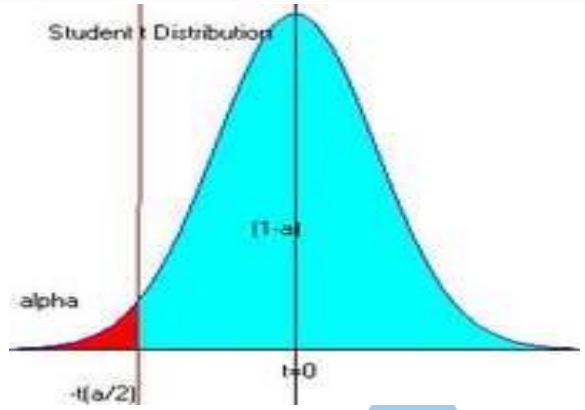
Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program

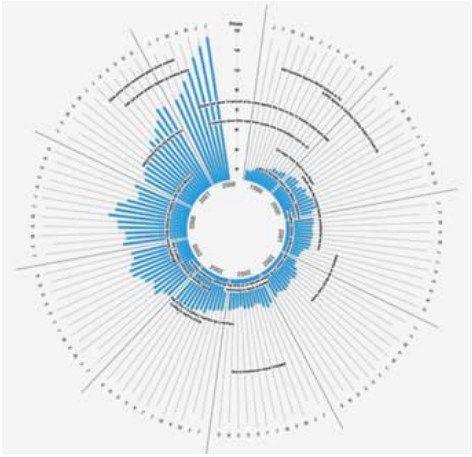


- Ada dua jenis hipotesis:
  - Hipotesis nol ($H_0$), yaitu klaim yang ingin ditunjukkan SALAH
  - Hipotesis satu ($H_1$), yaitu tandingan atau alternatif dari $H_0$
- Untuk kasus tukang ramal tadi:
  - $H_0$: $p = 1/3$ → menebak acak (ingin ditunjukkan bahwa ini salah)
  - $H_1$: $p > 1/3$ → punya kemampuan meramal
- Biasanya $H_0$ satu nilai atau kisaran nilai, tetapi $H_1$ merupakan kisaran nilai.
- Ada dua jenis hipotesis alternatif:
  - Bisa $H_1$: $p < 1/3$ (**satu arah**)
  - atau $H_1$: $p \neq 1/3$ (**dua arah**)



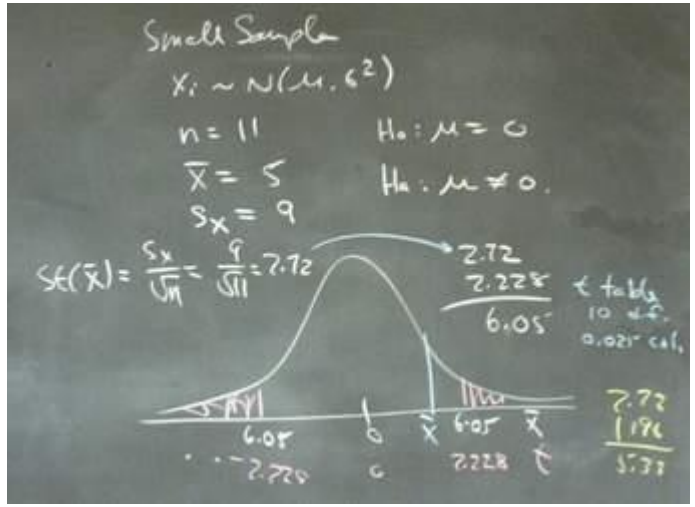

Boleh juga spt ini:
$H_0$: $p \leq 1/3$
$H_1$: $p > 1/3$

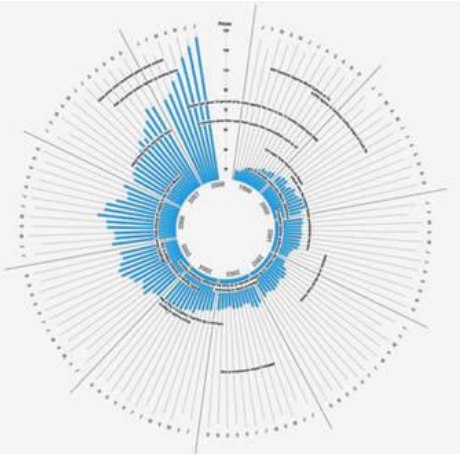


## Statistik uji dan nilai-p (*p-value*)

- Kembali ke tukang ramal, misal ternyata dia dapat menebak 40 orang dgn benar (sesuai data kepribadian):
  - Jadi statistik $\hat{p} = 40/116 = 0.345$ → ini dugaan bagi parameter $p$
- Hipotesis yang akan diuji:
  - $H_0$: $p = 1/3$ → menebak acak
  - $H_1$: $p > 1/3$ → punya kemampuan meramal
- Statistik $\hat{p}$ dibandingkan dengan parameter $p = 1/3$ menggunakan uji $z$ (uji berdasarkan sebaran Normal).
- Nilai-$p$ (*p-value*) mrp ukuran seberapa mungkin kita memperoleh statistik $\hat{p}$ *sebesar* yang ada ditangan kita jika seandainya $H_0$ benar. → jika nilai-p kecil, maka $H_0$ ditolak.

Boleh juga spt ini:
$H_0$: $p \leq 1/3$
$H_1$: $p > 1/3$



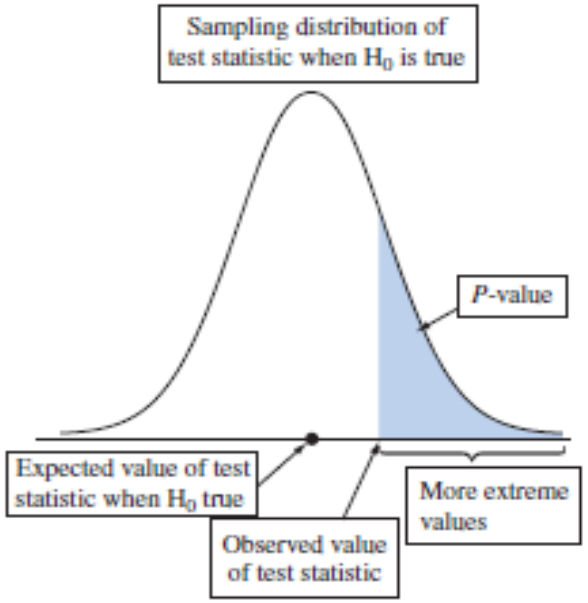



## Nilai-p (*p-value*)

- Nilai-p dihitung dari statistik uji yang digunakan;
- Untuk kasus ini statistik ujinya adalah:

$$Z_{hit} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- *Nilai-p* = $P(z > Z_{hit} \mid p_0)$, dengan $p_0$ adalah nilai yang ada pada hipotesis nol. Jika $H_0$: $p = \frac{1}{3}$ maka $p_0 = \frac{1}{3}$.

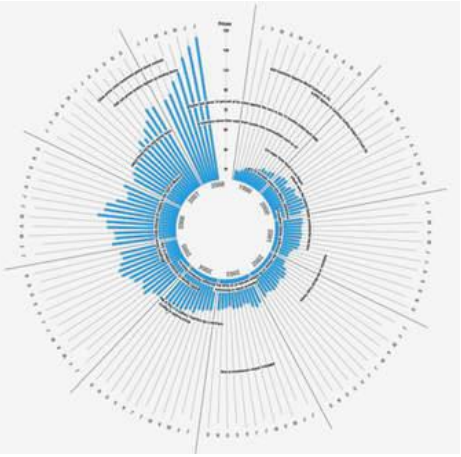



- Misal tukang ramal tadi dapat menebak 40 orang dengan benar (yaitu tebakan sesuai dengan data kepribadian), bagaimana uji hipotesisnya?

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program

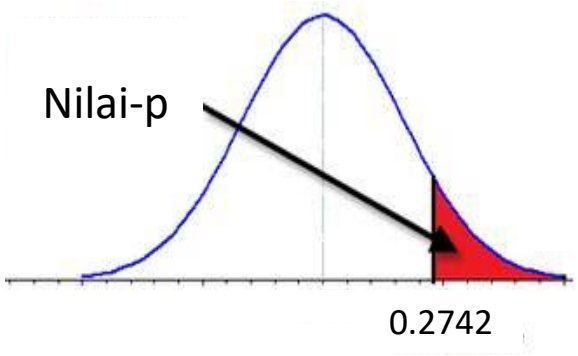


- Tukang ramal menebak benar 40 orang:
  - Statistik $\hat{p} = 40/116 = 0.345$ → penduga parameter $p$
- Hipotesis yang akan diuji:

  $H_0$: $p = 1/3$ → menebak acak

  $H_1$: $p > 1/3$ → punya kemampuan meramal
- Statistik ujinya

$$Z_{hit} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.345 - 0.333}{\sqrt{\frac{0.333 * 0.667}{116}}} = 0.2742$$

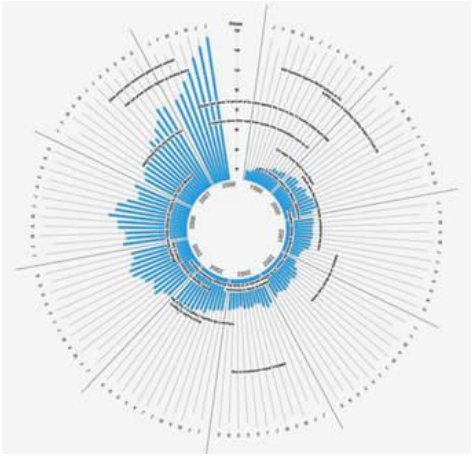



- Nilai-$p$ = $P(z > 0.2742 | p = \frac{1}{3}) = 0.39$ (lihat tabel normal)
- Karena nilai-$p$ besar (lebih dari 5%) maka data tdk berpihak pada $H_1$ sehingga kita MENERIMA $H_0$.

Boleh juga spt ini:

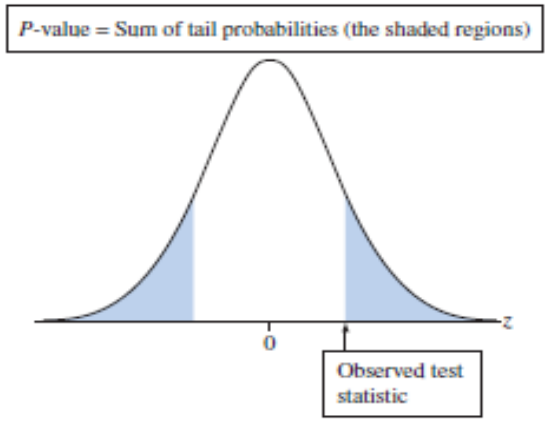
$H_0$: $p \leq 1/3$

$H_1$: $p > 1/3$



## Uji dua arah (*two-way*)

- Kita telah melakukan uji satu arah (*one-way*):
  - $H_0$: $p = 1/3$ → menebak acak
  - $H_1$: $p > 1/3$ → punya kemampuan meramal
- Bgm jika kita ingin melakukan uji dua arah (*two-way*):
  - $H_0$: $p = p_0$ vs $H_1$: $p \neq p_0$
- Statistik uji: $Z_{hit} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$
- Jika nilai-$p = P(|z| > Z_{hit} \mid p_0)$ *kecil* → menolak $H_0$
- Untuk tukang ramal: $Z_{hit} = \dfrac{0.345 - 0.333}{\sqrt{\dfrac{0.333 * 0.667}{116}}} = 0.2742$
- Nilai-$p = P(|z| > 0.2742 \mid p = \frac{1}{3}) = 0.78$ (tabel normal)
- Jadi menerima $H_0$: $p = 1/3$, artinya belum cukup bukti untuk mengatakan bhw peramal tidak menebak secara acak.




# Diskusi Dulu.....

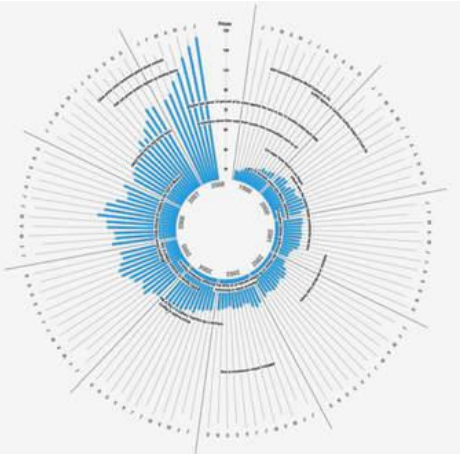
Sesi 1… beres!!!

Sesi 2…

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



Regardless of age, about 20% of American adults participate in fitness activities at least twice a week. However, these fitness activities change as the people get older, and occasionally participants become nonparticipants as they age. In a local survey of $n = 100$ adults over 40 years old, a total of 15 people indicated that they participated in a fitness activity at least twice a week. Do these data indicate that the participation rate for adults over 40 years of age is significantly less than the 20% figure? Calculate the $p$-value and use it to draw the appropriate conclusions.


Di Amerika sekitar 20% orang dewasa melakukan *fitness*. Dari survei terhadap 100 orang umur > 40 th ada 15 orang yang melakukan *fitness* sedikitnya dua kali seminggu. Bisakah dipercaya bahwa partisipasi orang berumur > 40 th dlm olahraga *fitness* lebih kecil dari 20%?

**Solution** Assuming that the sampling procedure satisfies the requirements of a binomial experiment, you can answer the question posed using a one-tailed test of hypothesis:

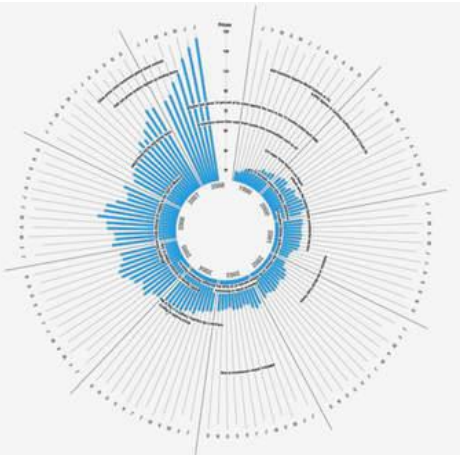
$H_0: p = 0.2$ vs $H_1: p < 0.2$ →

Boleh juga spt ini:
$H_0: p \geq 0.2$
$H_1: p < 0.2$

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



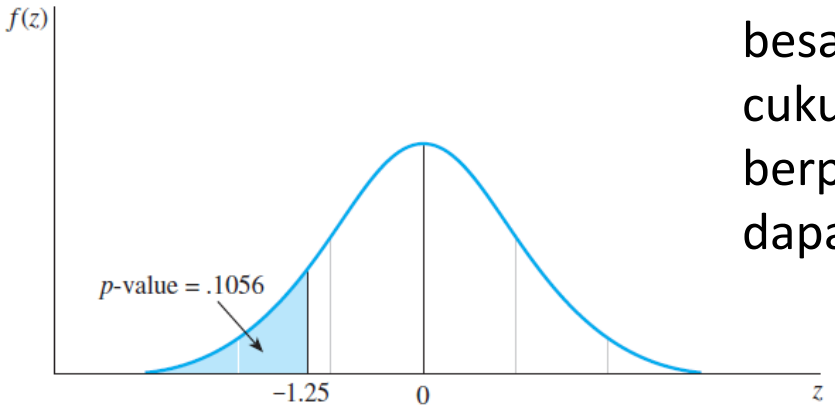
The observed value of $\hat{p}$ is 15/100 = .15 and the test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{.15 - .20}{\sqrt{\frac{(.20)(.80)}{100}}} = -1.25$$

← Ambil nilai $p_0$ = 0.2 sesuai dgn $H_0$

The $p$-value associated with this test is found as the area under the standard normal curve to the left of $z = -1.25$ as shown in Figure 9.10. Therefore,

$$p\text{-value} = P(z < -1.25) = .1056$$



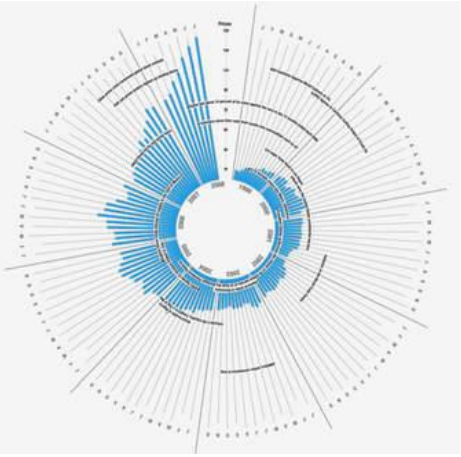

Ternyata nilai-$p$ = 0.1056 ini cukup besar (lebih besar 0.05) sehingga $H_0$ diterima, artinya tidak cukup bukti bahwa 20% rakyat Amerika berpartisipasi dalam fitness. Klaim tersebut tidak dapat disalahkan.

Coba pikirkan seandainya nilai-$p$ = 0.001, akankah kita menolak kebenaran $H_0$ ???

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



## Langkah-langkah pengujian hipotesis:

SUMMARY: Steps of a Significance Test for a Population Proportion $p$

1. **Assumptions**
   - Categorical variable, with population proportion $p$ defined in context.
   - Randomization, such as a simple random sample or a randomized experiment, for gathering data.
   - $n$ large enough to expect at least 15 successes and 15 failures under $H_0$ (that is $np_0 \geq 15$ and $n(1 - p_0) \geq 15$). This is mainly important for one-sided tests.
2. **Hypotheses**

   *Null:* $H_0$: $p = p_0$, where $p_0$ is the hypothesized value.

   *Alternative:* $H_a$: $p \neq p_0$ (two-sided) or $H_a$: $p < p_0$ (one-sided) or $H_a$: $p > p_0$ (one-sided)
3. **Test statistic**

$$z = \frac{\hat{p} - p_0}{se_0} \text{ with } se_0 = \sqrt{p_0(1 - p_0)/n}$$

4. **P-value**

| Alternative hypothesis | P-value |
|---|---|
| $H_a$: $p > p_0$ | Right-tail probability |
| $H_a$: $p < p_0$ | Left-tail probability |
| $H_a$: $p \neq p_0$ | Two-tail probability |

5. **Conclusion**

   Smaller P-values give stronger evidence against $H_0$. If a decision is needed, reject $H_0$ if the P-value is less than or equal to the preselected significance level (such as 0.05). Relate the conclusion to the context of the study.

Asumsi: data berasal dari contoh acak, $n$ cukup besar.

Hipotesis:
$H_0$ dan $H_1$ apakah satu arah atau dua arah.

Statistik uji:

$$Z_{hit} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Nilai-$p$:

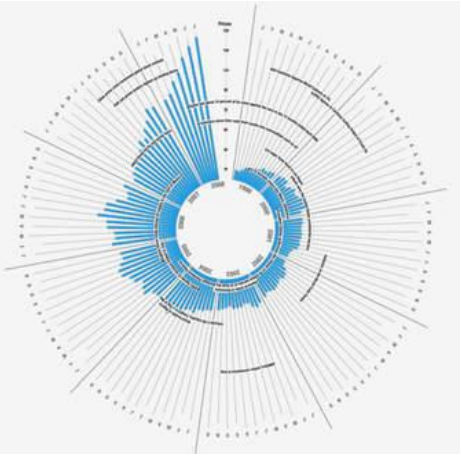
- Satu arah $P(Z > Z_{hit} | H_0 \text{ benar})$
- Dua arah $P(|Z| > Z_{hit} | H_0 \text{ benar})$

Kesimpulan: jika nilai-$p$ <0.05 maka $H_0$ ditolak

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



## Pengujian parameter nilaitengah $\mu$

- Telah dibahas cara menguji parameter proporsi ($p$).
- Bagaimana kalau yg ingin diuji adalah nilaitengah ($\mu$)?
- Prinsipnya sama dgn menguji parameter proporsi.

Peubah respon: durasi kerja (skala kontinu).

Ingin diketahui nilai-tengah durasi kerja ($\mu$) apakah 40 jam/minggu atau tidak.

Jadi ingin diuji :

$H_0$: $\mu = 40$ vs $H_1$: $\mu \neq 40$ (uji dua arah).

## The 40-Hour Work Week

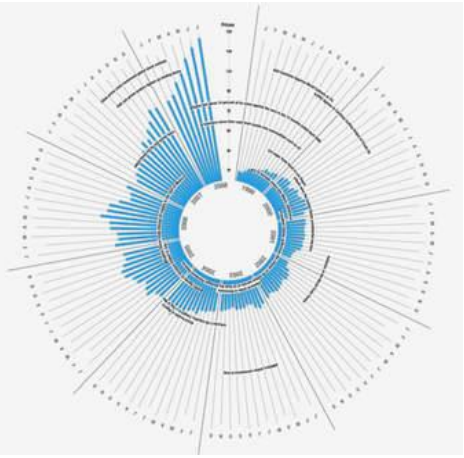
### Picture the Scenario

Since the Fair Labor Standards Act was passed in 1938, the standard work week in the United States has been 40 hours. In recent years, the standard work week has fallen to less than 40 hours in most of Western Europe and in Australia. But many believe that the work-oriented culture in the United States has resulted in pressure among workers to put in longer hours than the 40-hour standard. In industries such as investment banking, a 40-hour work week is considered slacker behavior and may result in losing a job. Because the mean number of working hours may differ between males and females (perhaps due to females working more part-time jobs), here we will analyze a sample of working hours for males. Exercise 9.32 conducts a similar analysis for a sample of females.

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program

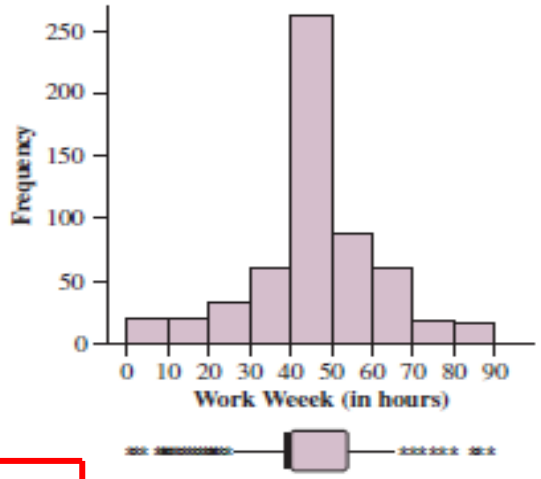


1. **Asumsi:**

   Peubah respon kuantitatif, bukan kategorik;

   Data dari contoh acak yang mendekati sebaran normal.

   For those who were working in 2012, the General Social Survey asked, "How many hours did you work last week?" For the 583 men included in the survey, the mean was 43.5 hours with a standard deviation of 15.3 hours.




2. **Hipotesis:**

   $H_0$: $\mu = \mu_0$ vs $H_1$: $\mu \neq \mu_0$ (dua arah)

   Mungkin juga $H_1$: $\mu > \mu_0$ atau $H_1$: $\mu < \mu_0$ (satu arah)

   Besarnya $\mu_0$ dalam kasus ini adalah 40 shg $H_0$ : $\mu = 40$.

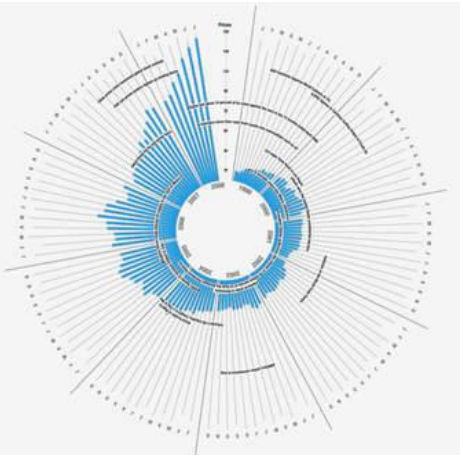
3. **Statistik uji:**

$$t = \frac{(\bar{x} - \mu_0)}{s_{\bar{x}}} = \frac{(\bar{x} - \mu_0)}{s/\sqrt{n}} = (43.5 - 40)/0.636 = 5.54.$$

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



***4. Nilai-p*:**

Nilai ini tergantung apakah uji satu arah atau dua arah;

Kasus yg kita bahas adalah uji dua arah shg nilai-p ada di ujung kiri dan ujung kanan.

**Normal baku (*z*)**

**5. Kesimpulan:**

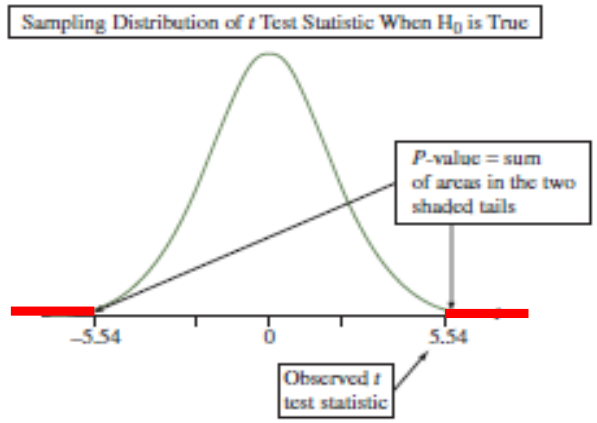
Karena n = 583 → P($|t_{(582)}|$>5.54) = kecil sekali, lebih kecil dari 0.001.

Jadi $H_0$ : $\mu = 40$ ditolak. Artinya ada cukup bukti bahwa durasi kerja dari para buruh itu bukan 40 jam/minggu, melainkan lebih dari 40 jam/minggu. Ingat statistik rata-rata durasi menunjukkan 43.5.

Selanjutnya SK 95% bagi $\mu$ adalah :

$$42.271<\mu<44.768$$

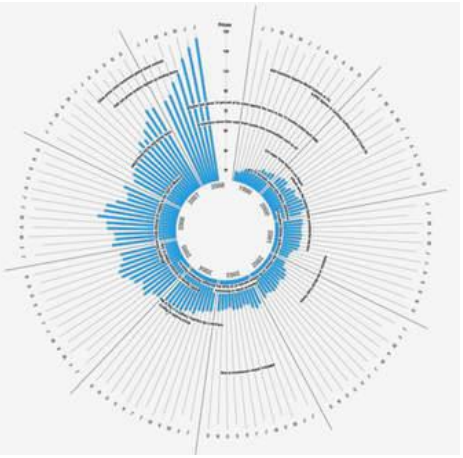
Jadi kita yakin 95% bahwa $\mu$ = 40.0 ada di luar selang tersebut.




# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



Apakah konsumsi garam utk orang AS telah berlebihan? Batas konsumsi 3300 mg/hr.

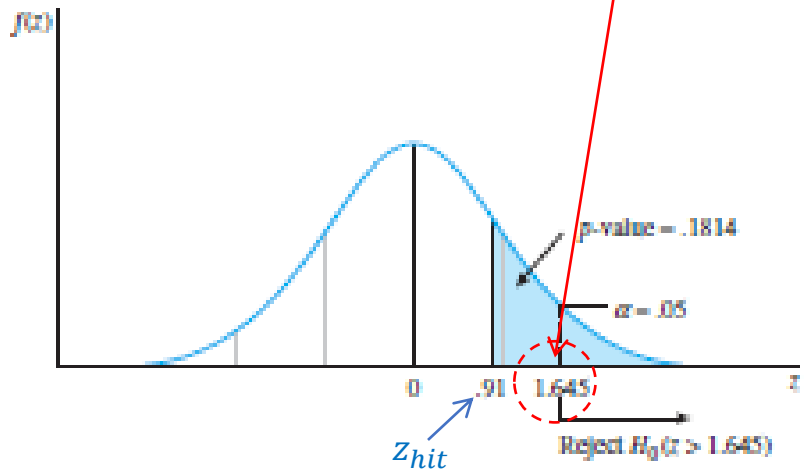
Hasil survei: rata-rata konsumsi 3400 mg dengan simpangan baku 1100 mg.

$H_0$: $\mu$ = 3300 vs $H_1$: $\mu$ > 3300.

Standards set by government agencies indicate that Americans should not exceed an average daily sodium intake of 3300 milligrams (mg). To find out whether Americans are exceeding this limit, a sample of 100 Americans is selected, and the mean and standard deviation of daily sodium intake are found to be 3400 mg and 1100 mg, respectively. Use $\alpha$ = .05 to conduct a test of hypothesis.

- **Statistik uji (*n=100*, gunakan *z*):** $z_{hit} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{3400-3300}{1100/\sqrt{100}} = 0.91$

**Menggunakan pendekatan nilai kritis:** untuk taraf nyata $\alpha$ = 0.05 uji satu arah memiliki nilai kritis $z_{.05}$ = 1.645. Karena $z_{hit}$ = 0.91 lebih besar dari $z_{.05}$ maka $H_o$ diterima. Artinya tidak cukup bukti bahwa konsumsi sodium orang Amerika sudah melebihi ambang batas. (INKONKLUSIF)




- ***Menggunakan pendekatan nilai-p:***
  - ***Nilai-p* = 0.1814 > 0.05 → terima $H_0$**
  - Statistik *z* = 0.91 < 1.645 ($z_{0.05}$) → terima $H_0$
  - Jadi hasil pengujian berdasar nilai-p dan berdasar nilai kritis sama saja.

# Diskusi Dulu.....

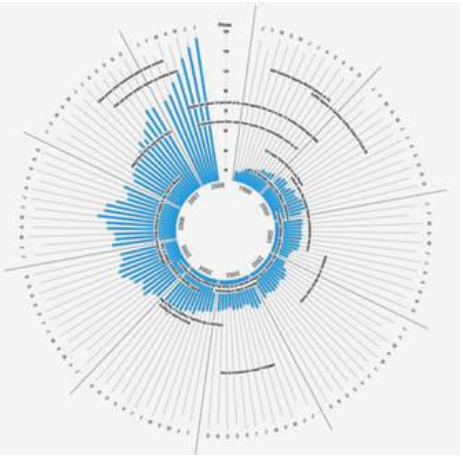
Sesi 1… beres!!!

Sesi 2…

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program

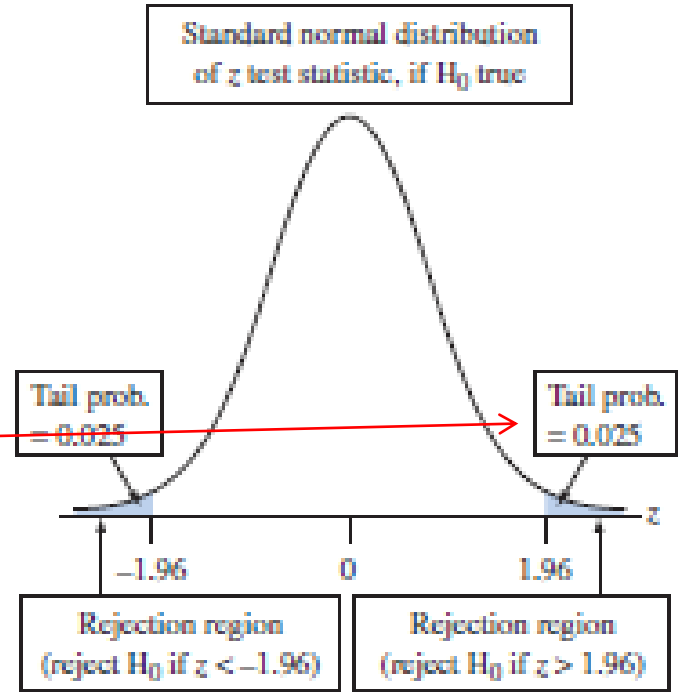


## Kesalahan dalam pengujian hipotesis

- Hasil pengujian hipotesis berpotensi untuk salah. Dalam hal ini ada 2 jenis kesalahan dalam uji hipotesis:
  - Ketika menolak $H_0$ pada hal $H_0$ benar → Salah Jenis I
  - Ketika menerima $H_0$ pada hal $H_0$ salah → Salah Jenis II

| Keputusan | Keadaan sesungguhnya | |
|---|---|---|
| | $H_0$ benar | $H_0$ salah |
| Tolak $H_0$ | *Jenis I* | ✓ |
| Terima $H_0$ | ✓ | *Jenis II* |

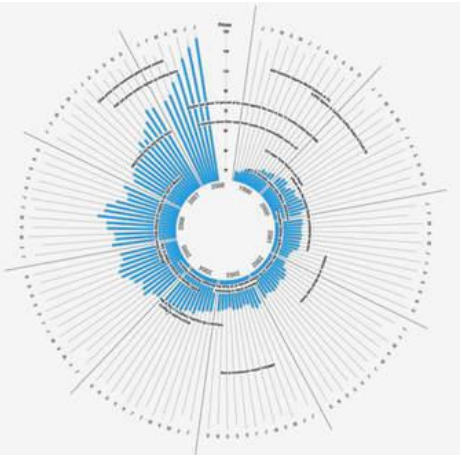
- **Peluang** salah jenis I → $\alpha$
- **Peluang** salah jenis II → $\beta$



# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



## Kuasa uji (*Power of Test*)

**Definisi:**

Kuasa uji adalah peluang menolak $H_0$ ketika $H_1$ benar. Kuasa uji = $(1-\beta)$ dan merupakan ukuran kemampuan uji tersebut berkinerja sesuai harapan.

| Keputusan | Keadaan sesungguhnya | |
|---|---|---|
| | $H_0$ benar | $H_0$ salah |
| Tolak $H_0$ | *Salah Jenis I* | Kuasa uji |
| Terima $H_0$ | ✓ | *Salah Jenis II* |

The daily yield for a local chemical plant has averaged 880 tons for the last several years. The quality control manager would like to know whether this average has changed in recent months. She randomly selects 50 days from the computer database and computes the average and standard deviation of the $n = 50$ yields as $\bar{x} = 871$ tons and $s = 21$ tons, respectively. Test the appropriate hypothesis using $\alpha = .05$.

Solution

**Null and alternative hypotheses:**

$H_0: \mu = 880$ versus $H_a: \mu \neq 880$

**Test statistic:** The point estimate for $\mu$ is $\bar{x}$. Therefore, the test statistic is

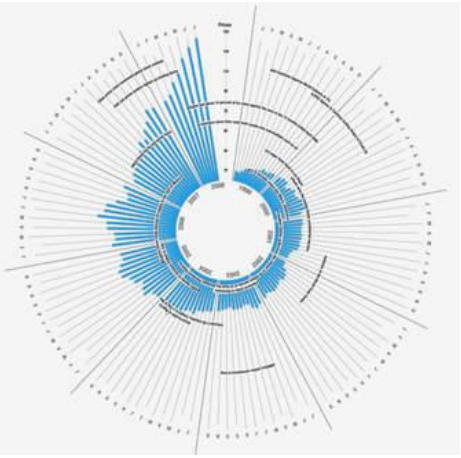
$$z \approx \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{871 - 880}{21/\sqrt{50}} = -3.03$$

**Kapasitas produksi selama ini adalah 880 ton. Ingin diketahui apakah ada perubahan kapasitas dlm bbrp bulan terakhir. Diambil contoh acak 50 hari dari database, ternyata diperoleh kapasitas rata-rata $\bar{x}$ = 871 ton dan $s$ = 21 ton. Ujilah ($\alpha$=0.05) apakah ada perubahan kapasitas produksi?**

# Pengujian Hipotesis

Hypothesis Testing

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



**Rejection region:** For this two-tailed test, you use values of $z$ in both the right and left tails of the standard normal distribution. Using $\alpha = .05$, the **critical values** separating the rejection and acceptance regions cut off areas of $\alpha/2 = .025$ in the right and left tails. These values are $z = \pm 1.96$ and the null hypothesis will be rejected if $z > 1.96$ or $z < -1.96$.

**Conclusion:** Since $z = -3.03$, the calculated value of $z$, falls in the rejection region, the manager can reject the null hypothesis that $\mu = 880$ tons and conclude that it has changed. The probability of rejecting $H_0$ when $H_0$ is true is $\alpha = .05$, a fairly small probability. Hence, she is reasonably confident that the decision is correct.

Berapa kuasa uji (*1- β*) jika nilai $\mu$ = 870 ton?

$$z_1 \approx \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{874.18 - 870}{21/\sqrt{50}} = 1.41$$

$$z_2 \approx \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{885.82 - 870}{21/\sqrt{50}} = 5.33$$

$\beta = P(\text{accept } H_0 \text{ when } \mu = 870) = P(874.18 < \bar{x} < 885.82 \text{ when } \mu = 870)$

$= P(1.41 < z <= 1 - .9207 = .0793 \longrightarrow 1 - \beta = 1 - .0793 = .9207$

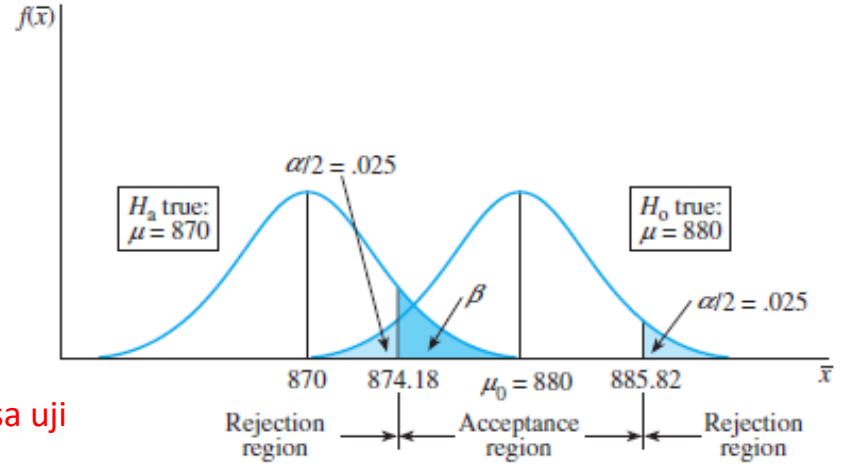
Kuasa uji

Daerah penolakan untuk α=0.05 adalah z>1.96 atau z<-1.96. Karena $z_{hit}$= -3.03 ada di daerah penolakan maka kita menolak $H_0$: $\mu$ = 880.

Kuasa uji (1-$\beta$):

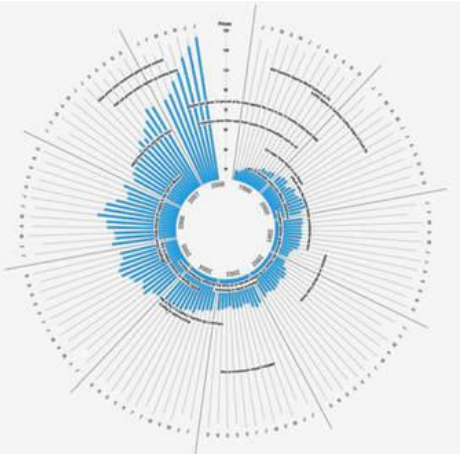
Besarnya $\beta$ sama dengan peluang menerima $H_0$: $\mu$ = 880, jika $H_0$ salah, misal yg benar $\mu$ = 870.

Lihat gambar: $\beta$ = 0.0793, shg kuasa uji adalah 0.9207.



# Ringkasan Pengujian Hipotesis

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



Misal $x_1, x_2, \ldots, x_n$ adalah contoh acak dari populasi normal dengan nilai-tengah $\mu$ dan ragam $\sigma^2$ sehingga diperoleh statistik $\bar{x}$ dan $s^2$. Selain itu $\hat{p}$ merupakan penduga parameter proporsi $p$. Pengujian hipotesis dapat berpedoman pada tabel berikut ini.

| Hipotesis | Ragam | Ukuran contoh | Statistik uji | Nilai-p |
|---|---|---|---|---|
| $H_0 : p = p_0$ | Boleh tidak diketahui | Besar, misal $n$ = 200 | $Z_{hit} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ | - Satu arah P($Z > Z_{hit} \mid H_0$ benar)<br>- Satu arah P($Z < Z_{hit} \mid H_0$ benar)<br>- Dua arah P($\lvert Z \rvert > Z_{hit} \mid H_0$ benar) |
| $H_0 : \mu = \mu_0$ | Diketahui $\sigma^2$ | Yang penting n > 2 | $Z_{hit} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ | - Satu arah P($Z > Z_{hit} \mid H_0$ benar)<br>- Satu arah P($Z < Z_{hit} \mid H_0$ benar)<br>- Dua arah P($\lvert Z \rvert > Z_{hit} \mid H_0$ benar) |
| $H_0 : \mu = \mu_0$ | Tidak diketahui | Besar, misal n > 30 | $Z_{hit} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ | - Satu arah P($Z > Z_{hit} \mid H_0$ benar)<br>- Satu arah P($Z < Z_{hit} \mid H_0$ benar)<br>- Dua arah P($\lvert Z \rvert > Z_{hit} \mid H_0$ benar) |
| $H_0 : \mu = \mu_0$ | Tidak diketahui | Kecil, misal n ≤ 30 | $t_{hit} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ , db=(n-1) | - Satu arah P($Z > t_{hit} \mid H_0$ benar)<br>- Satu arah P($Z < t_{hit} \mid H_0$ benar)<br>- Dua arah P($\lvert Z \rvert > t_{hit} \mid H_0$ benar) |

# Peran Komputer

## The Bootstrap: Using simulation to test a hypotesis.


- **Bisakah kita menguji hipotesis untuk suatu parameter yang sebaran statistiknya tidak diketahui?**
- **Masih ingat pengukuran untuk menilai konsistensi (baca simpangan baku) dari timbangan?**
- **Misal dari hasil pengukuran bobot badan (dalam *pound*) ingin diuji simpangan baku populasi $\sigma$. Maka kita kesulitan mendapatkan sebaran dari statistik *s* (*sampling distribution*).**
- **Efron (1979) memperkenalkan metode *bootstrap* untuk menguji galat baku ini (dan juga untuk parameter lain yang sebaran statistikmya tidak diperoleh secara matematika).**
- **Caranya adalah menyusun SK $(1-\alpha)100\%$ dari parameter $\sigma$ melalui simulasi ribuan kali sehingga SK tersebut bisa diperoleh berdasar hasil statistik *s* yang diperoleh pada setiap simulasi.**
- **Jika SK tersebut TIDAK mencakup nilai $\sigma_0$ maka hipotesis $H_0$: $\sigma = \sigma_0$ ditolak.**

# Peran Komputer

## The Bootstrap: Using simulation to test a hypotesis.

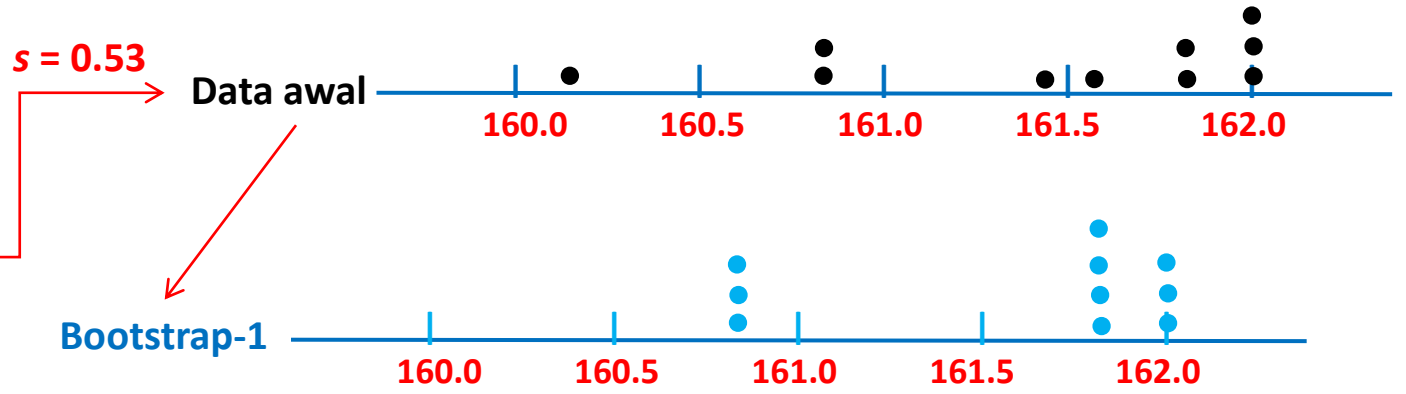
- Perhatikan persoalan menyusun selang kepercayaan dari simpangan baku populasi seperti berikut ini:
  - Ingin diketahui seberapa jauh variasi hasil pengukuran suatu timbangan.
  - Seseorang menimbang dirinya 10 kali, hasilnya sbb (pound):

    160.2, 160.8, 161.4, 162.0, 160.8, 162.0, 162.0, 161.8, 161.6, 161.8

- Randomly sample 10 observations from this sample data distribution. This corresponds to randomly sampling with replacement 10 of the original 10 observations (see the bubble labeled Resample 1). We did this and got

  162.0, 160.8, 161.8, 161.8, 162.0, 161.8, 161.8, 162.0, 160.8, 160.8

  For these 10 observations, the sample standard deviation is $s = 0.53$.

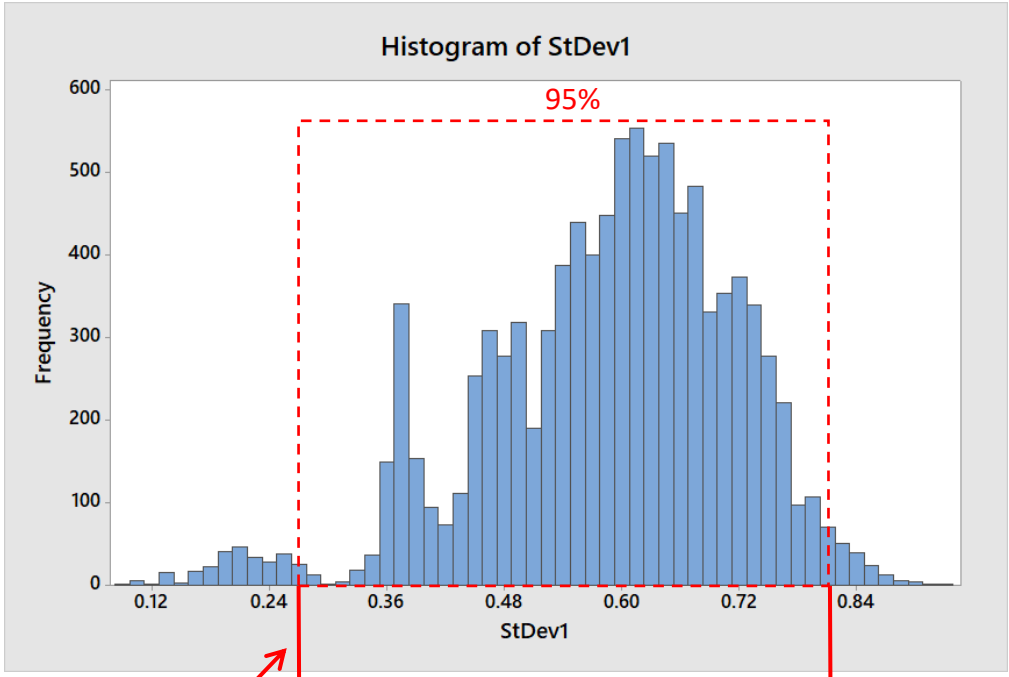


# Peran Komputer

## The Bootstrap: Using simulation to test a hypotesis.

- Repeat the preceding step 100,000 times, each time obtaining a resample of size 10 and computing $s$. This gives us 100,000 values of the sample standard deviation. Figure 8.11 shows a histogram of their
- Now identify the middle 95% of these 100,000 sample standard deviation values. For the 100,000 samples we took, the 2.5th percentile was 0.26 and the 97.5th percentile was 0.80. In other words, 95% of the resamples had sample standard deviation values between 0.26 and 0.80.

| Contoh | Bootstrap samples | $S$ |
|---|---|---|
| 1 | 160.8, 162.0, …, 160.2 | 0.53 |
| 2 | 162.0, 161.8, …, 161.6 | 0.69 |
| : | : | : |
| 100,000 | 162.0, 160.2, …, 161.4 | 0.36 |

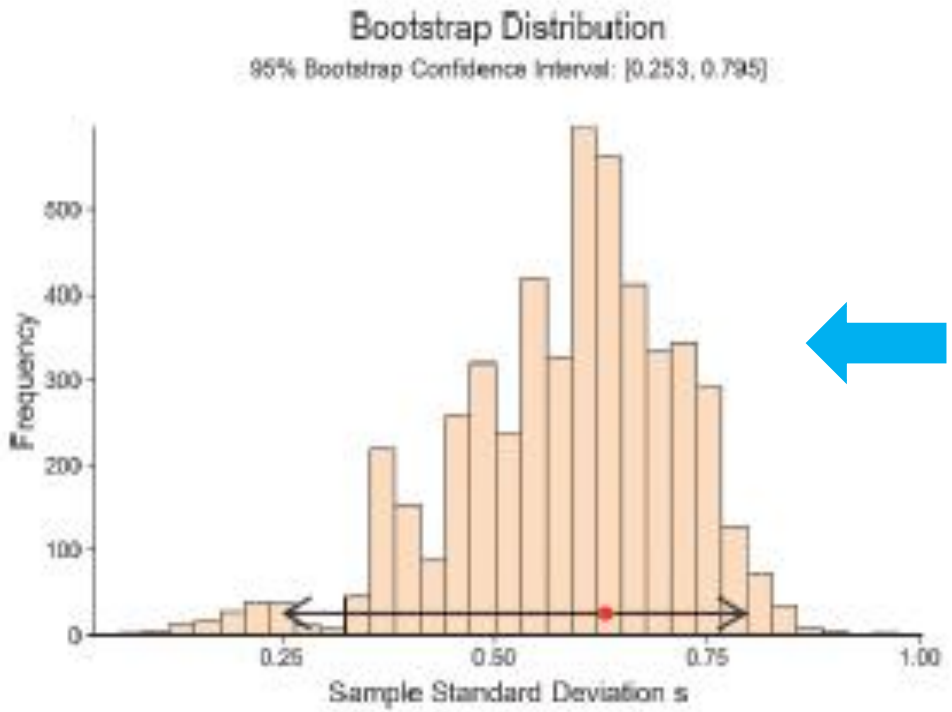



(0.2573, 0.7933)

Selang kepercayaan bootstrap utk $\sigma$.

# Peran Komputer

## The Bootstrap: Using Simulation to Construct a Confidence Interval



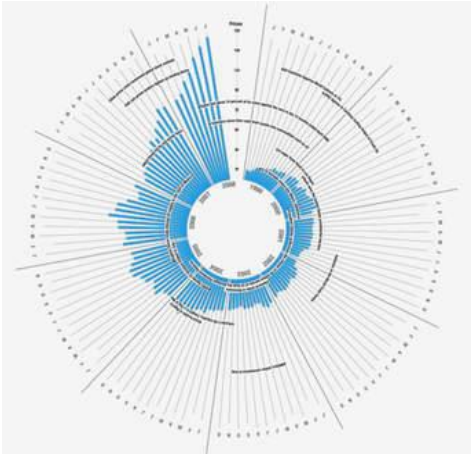

▲ **Figure 8.11** A Bootstrap Frequency Distribution of Standard Deviation Values. These were obtained by taking 100,000 samples of size 10 each from the sample data distribution. **Questions** Does the sampling distribution of the sample standard deviation look approximately normal? What is the practical reason for using the bootstrap method?

- Sebaran dari *s* nampak tidak simetrik, menjulur ke kiri.
- Penggunaan sebaran normal dalam menyusun SK menjadi tidak tepat.
- SK yg disusun adalah yang terpendek dari berbagai kemungkinan yg ada.
- Teknik bootstrap menjadi cukup sederhana dan menghasilkan SK utk $\sigma$, yaitu (**0.2573, 0.7933**)

Karena nilai 0 tidak dicakup oleh selang kepercayaan tsb maka hipotesis $H_0$: $\sigma = 0$ ditolak. Artinya timbangan tidak 100 persen akurat.

# Latihan

IPB University
Bogor Indonesia
Department of Statistics
Statistics and Data Science Study Program



## Mendenhall

1. **Halaman 339 soal nomor 9.6**
2. Halaman 340 soal nomor 9.8, 9.13
3. Halaman 350 soal nomor 9.35
4. Halaman 374 soal nomor 10.9
5. Halaman 375 soal nomor 10.15

**Diskusikan**

## Agresti

6. **Halaman 417 soal nomor 9.1, 9.2**
7. Halaman 431 soal nomor 9.12
8. Halaman 432 soal nomor 9.20
9. Halaman 443 soal nomor 9.34
10. Halaman 455 soal nomor 9.54

# Diskusikan.....

**Data pendapatan per bulan dari 30 petani (contoh acak) di desa Tegal Gundil.**
**Ujilah apakah mediannya sama dengan Rp 3.500.000 per bulan?**

**598236, 559348, 1393754, 2904429, 4172863, 1470580**
**3544143, 4918026, 558184, 1436698, 2598927, 2997189**
**6806786, 4623161, 830000, 3566447, 2707305, 3649276**
**2767137, 4233366, 2794188, 3116484, 4082507, 613158**
**7375163, 5112766, 1870627, 121805, 1330808, 863539**

**Hipotesis yg diuji adalah Ho: Med = 3.500.000**
**H1: Med tdk sama 3.500.000**

Urutan Bootstrap (umum) :

1. Ada data sebanyak **n** (contoh acak)
2. Ingin menguji/SK dari parameter $\theta$, tapi statistik $\hat{\theta}$ tidak diketahui sebarannya.

3. Lakukan Bootstrap:

- Resampling sebanyak **n** pengamatan terhadap **dengan pemulihan**
- Hitung statistik $\hat{\theta}$
- **Ulangi 100000 kali (**lakukan simulasi 100000 kali**)**shg diperoleh statistik $\hat{\theta}$ sebanyak 100000.
- Buat histogramnya dan tentukan batas kiri-kanan 95%
- Maka SK 95% diperoleh dan pengujian hipotesis bisa dilakukan.

4. Aplikasinya bisa menggunakan minitab (mudah)

# ☺ THANK YOU ☺

IPB University
Bogor Indonesia

Department of Statistics

Study Programs in Statistics and Data Science