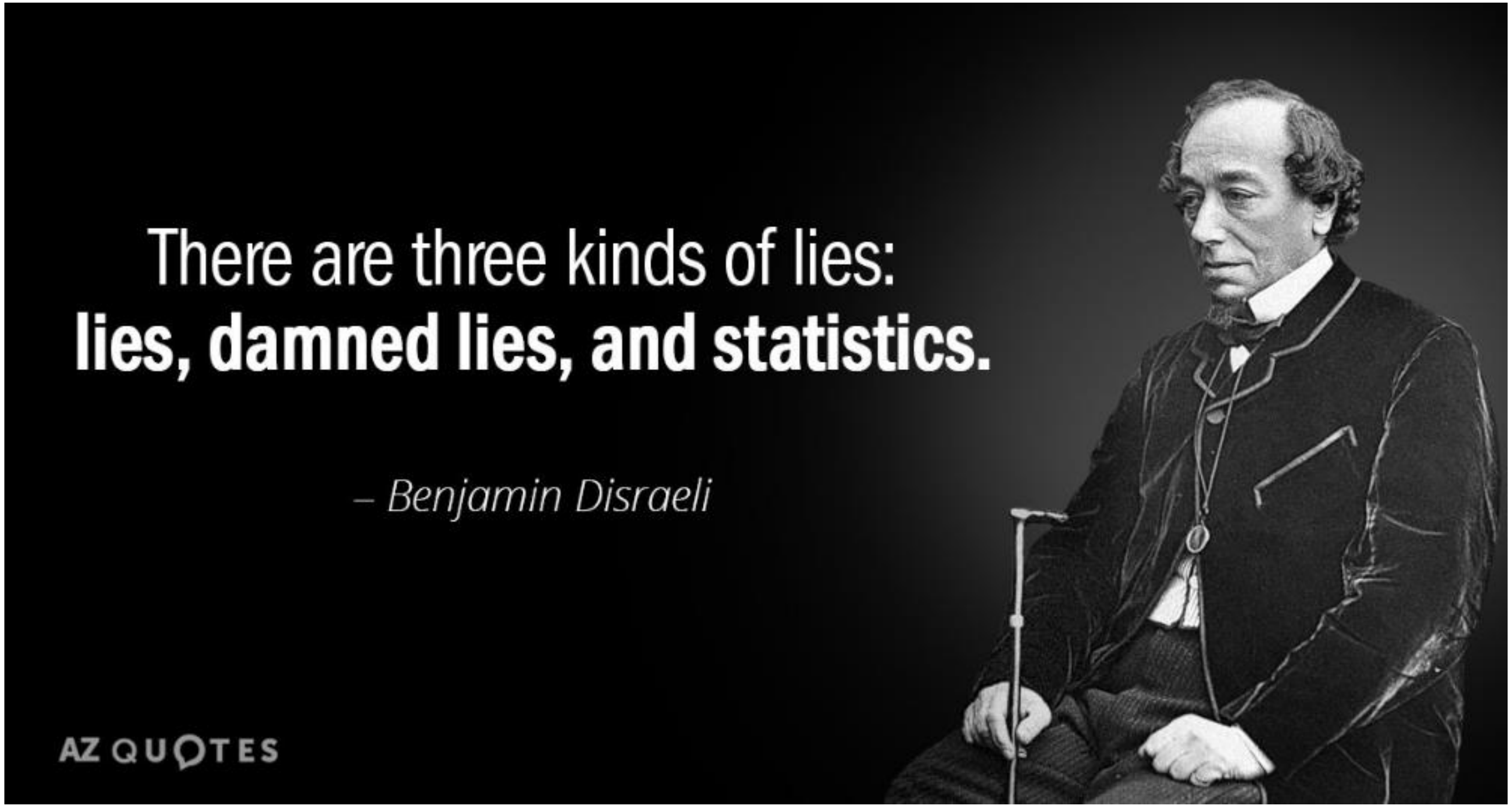
Statistika Non Parametrik
TSD - Ganjil 2023/2024

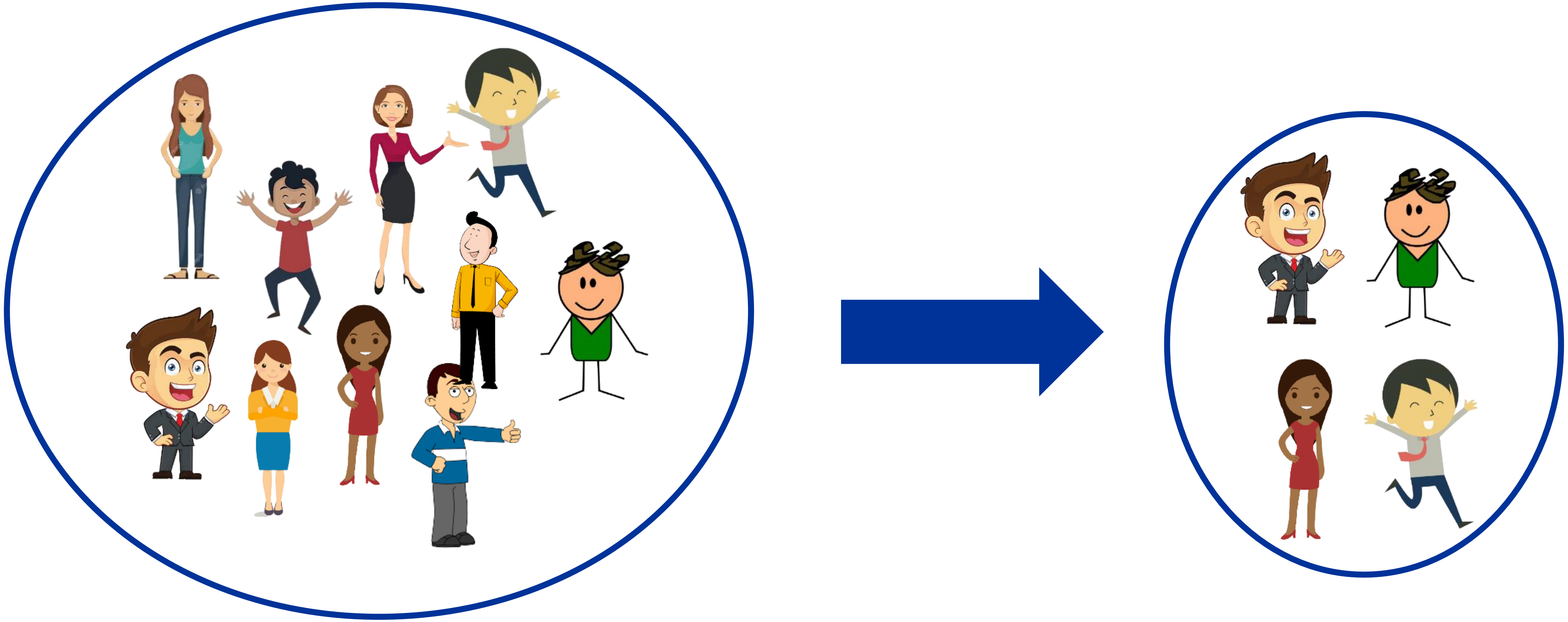
Pertemuan 2:
Konsep Dasar Statistika Non Parametrik

Review


There are three kinds of lies:
lies, damned lies, and statistics.
– Benjamin Disraeli
AZ QUOTES

Review

Review



POPULASI

↓

PARAMETER

SAMPEL

↓

STATISTIK

Review

Parameter dan Statistik

- Parameter adalah ukuran numerik dari suatu populasi.
- Statistik adalah ukuran numerik dari sampel.
- Parameter diestimasi/ diperkirakan dengan statistik.
- Statistik disebut juga dengan estimator dari parameter. Saat satu nilai digunakan sebagai estimasi, estimasi tersebut disebut estimasi titik dari populasi parameter.

Estimator (Sample Statistic)		Population Parameter
$\overline{X}$	$\xrightarrow{\text{estimates}}$	μ
S^2	$\xrightarrow{\text{estimates}}$	σ^2
$\hat{p}$	$\xrightarrow{\text{estimates}}$	p

Review

- Skala Pengukuran Data

	Skala Nominal	Skala Ordinal	Skala Interval	Skala Rasio
Mampu Membedakan	√	√	√	√
Memiliki Urutan		√	√	√
Memiliki Interval			√	√
Memiliki Nol Mutlak				√

0°, 50°,…

Peringkat

45kg, 62kg, …

- ✓ Statistical procedures:
  - Parametric (or “regular”) statistics procedures, and
  - **Nonparametric statistical procedures**

- ✓Nonparametric statistics pertama kali dikenalkan tahun 1942 oleh Wolfowitz.
- ✓Nonparametric statistics banyak digunakan dalam bidang kesehatan dan dalam sosial karena ketersediaan data yang terbatas lebih banyak berbentuk kategorik.

- ✓Nonparametric statistics digunakan lebih karena asumsi-asumsi dari pada analisis secara statistik parametrik tidak terpenuhi.
- ✓Nonparametric statistics juga dikenal sebagai prosedur bebas sebaran (*free distribution*).

Nonparametric Statistics

VS

- Tidak mensyaratkan distribusi atau asumsi
- Ukuran sampel kecil ($n < 30$)
- Digunakan pada skala data nominal atau ordinal

Parametric Statistics

- Mensyaratkan distribusi atau asumsi
- Ukuran sampel besar ($n > 30$)
- Digunakan pada skala data interval atau rasio

- Analisis Nonparametric statistics adalah bias karena analisis dilakukan pada data berupa peringkat (skala ordinal) ataupun tanda (skala nominal).
- Walaupun data pengamatan berskala interval atau rasio maka apabila analisis nonparametric dilakukan terhadap data tersebut, maka data yang berskala interval atau rasio haruslah dibuat dalam bentuk tanda atau peringkat, bergantung pada analisis apa yang akan digunakan.

Keuntungan Nonparametric Statistics

- Asumsinya lebih sedikit (Dibandingkan parametric statistics)
- Dapat digunakan pada lebih banyak bidang (kesehatan, social, dll)
- Bisa digunakan untuk sampel kecil
- Mudah digunakan (proses penghitungan lebih sederhana)
- Lebih mudah dipahami
- Tidak dipengaruhi oleh data ekstrim(outliers)
- Dapat digunakan pada skala data nominal dan ordinal. Sedangkan parametric cenderung diterapkan pada skala interval dan rasio.

Kekurangan Nonparametric Statistics

- Kurang powerful bila dibandingkan dengan parametric statistics, karena hanya memerlukan asumsi yang lebih sedikit.
- Tidak dapat digunakan untuk mengestimasi populasi (karena proses penghitungan yang sederhana, sampel kecil, dan tidak mensyaratkan distribusi normal).
- Mendorong orang untuk menggunakan Nonparametric statistics daripada parametric statistics walaupun jumlah sampel nya sama.
- Perlu usaha lebih untuk mengubah skala data menjadi nominal atau ordinal apabila data awal berskala interval atau rasio.

FAKULTAS TEKNOLOGI MAJU DAN MULTIDISIPLIN

Penggunaan Nonparametric Statistics

- Uji satu sampel
  - ✓Uji Run
  - ✓Uji Sign
  - ✓Uji Binomial
  - ✓Uji Wilcoxon

- Uji dua sampel independen
  - ✓Uji Median
  - ✓Uji Mann-Whitney
  - ✓Fisher Exact Probability
- Uji dua sampel dependen
  - ✓Uji Sign
  - ✓Mc Nemar
  - ✓Uji Wilcoxon 2 sampel

- Uji k-sampel independen
  - ✓Uji Kruskal Wallis
  - ✓Uji Perbandingan ganda
  - ✓Uji John Khere Terpstra
- Uji k-sampel dependen
  - ✓Uji Friedman
  - ✓Uji Perbandingan ganda
  - ✓Uji Page
  - ✓Uji Durbin

Penggunaan Nonparametric Statistics

- Uji kesesuaian distribusi
  - ✓Uji Chi-square
  - ✓Uji Kolmogorov-Smirnov
  - ✓Uji Liliefors

- Uji Asosiatif
  - ✓Korelasi Rank Spearman
  - ✓Korelasi Tau Kendall
  - ✓Kontingensi

Terima Kasih

Statistika Non Parametrik
TSD - Ganjil 2023/2024

Minggu Ke-2:
"Run Test dan Sign Test"

Pengujian Hipotesis 1 Sampel

- Pengujian hipotesis satu sampel untuk menguji perbedaan rata-rata sampel (observasi) dengan rata-rata yang diharapkan (populasi).
- **Tujuan**:
  - ✓Menguji perbedaan *central tendency* (lokasi) antara sampel dan populasi.
  - ✓Menguji perbedaan antara frekuensi observasi dengan frekuensi yang diharapkan.
  - ✓Menguji perbedaan antara proporsi observasi dengan proporsi yang diharapkan.
  - ✓Menguji apakah sampel diambil dari populasi dengan bentuk distribusi tertentu.
  - ✓Menguji apakah sampel diambil secara random dari populasi yang ada

Pedoman Memilih Teknik Statistik Non Parametrik Pengujian Hipotesis Satu Sampel

Skala yang Digunakan	Alat Analisis Pengujian Hipotesis Satu Sampel
Nominal	Uji Run Uji Sign Uji Binomial
Ordinal	Uji Wilcoxon Signed Rank

Run Test

Run Test

- Uji ini dimaksudkan untuk menguji keacakan (kerandoman) data dari suatu sampel.
- Uji ini didasarkan pada adanya runtun.
- Runtun adalah huruf-huruf atau tanda-tanda yang identik yang diikuti oleh satu huruf atau satu tanda yang berbeda.

® ® ® © © © ® © © © © ® ® © ®

1 2 3 4 5 6 7

Run Test

Contoh :

- ✓ aa bbb a b aa bb ada 6 runtun
- ✓ ++++ -- + - + - + - ada 8 runtun
- ✓ Dua sampel; sampel I dan sampel II sebagai berikut :

  Sampel I : 5, 16, 12, 17, 8, 9, 12

  Sampel II : 20, 7, 14, 19, 10

  Jika kedua sampel digabungkan dan datanya dituliskan/ disusun secara terurut dari kecil kebesar, maka : 5, 7, 8, 9, 10, 12, 12, 14, 16, 17, 19, 20

  Deretan diatas dapat pula dituliskan sebagai berikut :

  I, II, I, I, II, I, II, II, I, I, II, II jadi ada 8 runtun

Run Test

- Jika ukuran sampel besar **(≥ 30)** maka distribusi mendekati distribusi normal, sehingga untuk pengujian hipotesis dilakukan konversi banyaknya jumlah Runs ke dalam nilai Z yang mengikuti distribusi normal.

$$\boxed{Z = \frac{r - \mu_r}{\sigma_r}}$$

Untuk daerah penolakan / keputusan, bandingkan nilai Zhitung dengan Ztabel (dari tabel distribusi normal)

Keterangan:

r = Banyaknya runs.

μ_r = Mean

σ_r = Deviasi standar

n_1 = Banyaknya sampel salah satu group.

n_2 = Banyaknya sampel group yang lain.

Run Test

- Mean = $\mu_r = \frac{2n_1n_2}{n_1+n_2} + 1$
- Deviasi Standard = $\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$

Run Test

Perumusan Hipotesis :

H_0 : Data pengamatan telah diambil secara acak dari suatu populasi

H_1 : Data pengamatan diambil dari populasi tidak acak

Atau

H_0 : Sampel yang diambil dari suatu populasi adalah acak

H_1 : Sampel yang diambil dari suatu populasi tidak acak

Run Test

- **Statistik Uji :**

  r = banyaknya runtun yang terjadi

- **Daerah Kritis :**

  Tolak H_0, bila : $r \leq r_{bawah}$ atau $r \geq r_{atas}$

  r_{bawah} dan r_{atas} dari Tabel nilai kritis untuk runtun r dengan n_1 dan n_2

  Dimana :

  n_1 : Banyak data bertanda (+) atau huruf tertentu

  n_2 : Banyak data bertanda (-) atau huruf lainnya

Run Test

- **Contoh 1**

Dari sebuah pengukuran pengetahuan tentang ASI eksklusif pada 18 orang ibu hamil, diperoleh skor median sebesar 72. Artinya, ibu hamil dengan skor ≥ 72 adalah ibu hamil dengan kategori pengetahuan baik dan sebaliknya, ibu hamil dengan skor < 72 adalah ibu hamil dengan kategori pengetahuan kurang baik. Bagaimanakah keputusan hipotesisnya dengan derajat kepercayaan 95 % dan derajat signifikansi 5 % ?

Run Test

- **Hipotesis**

H0 = Tidak ada perbedaan pengetahuan ibu hamil. Hal ini berarti urutan dalam memiliki pengetahuan bersifat random

H1 = Ada perbedaan pengetahuan ibu hamil. Hal ini berarti urutan dalam memiliki pengetahuan tidak bersifat random

Run Test

- **Data pengamatan**

- Buat koding:

Coding 0 : >= 72 pengetahuan baik

Coding 1 : < 72 pengetahuan kurang

n1 : coding 0 = 10

n2 : coding 1 = 8

ID	Skor	Coding
1	65	1
2	32	1
3	87	0
4	96	0
5	88	0
6	54	1
7	52	1
8	48	1
9	67	1
10	78	0
11	43	1
12	56	1
13	78	0
14	94	0
15	84	0
16	85	0
17	92	0
18	76	0

Upper critical values of r in the runs test

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

Source: Frieda S. Swed and C. Eisenhart, "Tables for Testing Randomness of Grouping in a Sequence of Alternatives," *Ann. Math. Statist.*, 14 (1943), 66–87.

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

Source: Frieda S. Swed and C. Eisenhart, "Tables for Testing Randomness of Grouping in a Sequence of Alternatives," *Ann. Math. Statist.*, 14 (1943), 66–87.

Note: For the one-sample runs test, any value of r that is equal to or smaller than that shown in the body of this table for given value of n_1 and n_2 is significant at the 0.05 level.

Run Test

- Jumlah run = 6
- Dari tabel runs test didapatkan
- r_{bawah} = 5
- r_{atas} = 15

- **Keputusan**:
- Karena $r > r_{bawah}$ atau $r < r_{atas}$ maka gagal tolak H0

- **Kesimpulan**:
- tidak ada perbedaan pengetahuan ibu hamil. Hal ini berarti urutan dalam memiliki pengetahuan bersifat random.

ID	Skor	Coding	
1	65	1	1
2	32	1	
3	87	0	2
4	96	0	
5	88	0	
6	54	1	3
7	52	1	
8	48	1	
9	67	1	
10	78	0	4
11	43	1	5
12	56	1	
13	78	0	6
14	94	0	
15	84	0	
16	85	0	
17	92	0	
18	76	0	

Run Test

- Hasil SPSS

Runs Test

	Nilai
Test Value[a]	72.00
Total Cases	18
Number of Runs	6
Z	-1.669
Asymp. Sig. (2-tailed)	.095

a. User-specified.

Gagal tolak H0

Tidak ada perbedaan pengetahuan ibu hamil. Hal ini berarti urutan dalam memiliki pengetahuan **bersifat random**.

Run Test

- **Contoh 2**

Pak lurah mendengar bahwa beberapa pemuda desa Suka Ribut suka merokok. Untuk menguji apakah pemuda desa Suka Ribut yang suka merokok bersifat acak atau tidak, diambil sampel 10 pemuda.

Run Test

- **Data Penelitian**

Nama pemuda	Sikap
Tuji	Tidak
Taryo	Suka
Giwan	Tidak
Marso	Suka
Prapto	Suka
Gito	Tidak
Sunar	Suka
Yono Antasena	Suka
Cipto	Suka
Paryan	Tidak

Run Test

- **Penyelesaian**

1. Judul Penelitian

   Analisis Kerandoman Selera Merokok di Desa Suka Ribut

2. Variabel Penelitian

   Selera terhadap Rokok

3. Pertanyaan Penelitian

   Apakah pemuda di Desa Suka Ribut yang suka merokok bersifat acak?

4. Hipotesis:

   H_o : Pemuda di Desa Suka Ribut yang suka merokok bersifat acak.

   H_a : Pemuda di Desa Suka Ribut yang suka merokok tidak bersifat acak

5. Kriteria Pengujian

   H_o tidak dapat ditolak, jika Sig. > alpha (α).

   H_a diterima, jika Sig. ≤ alpha (α).

Run Test

- **Analisis Data**

Karena untuk menguji kerandoman data untuk satu sampel, maka digunakan Run Test.

<table>
<tr><th>No</th><th>Selera</th><th>Runs</th><th>No</th><th>Sikap</th><th>Runs</th></tr>
<tr><td>1.</td><td>T</td><td>1</td><td>6.</td><td>T</td><td>5</td></tr>
<tr><td>2.</td><td>S</td><td>2</td><td>7.</td><td>S</td><td rowspan="3">6</td></tr>
<tr><td>3.</td><td>T</td><td>3</td><td>8.</td><td>S</td></tr>
<tr><td>4.</td><td>S</td><td rowspan="2">4</td><td>9.</td><td>S</td></tr>
<tr><td>5.</td><td>S</td><td>10.</td><td>T</td><td>7</td></tr>
<tr><td>Jumlah Runs</td><td colspan="4"></td><td>7</td></tr>
</table>

Run Test

- **Output SPSS**

Runs Test

	Sikap Thp Rokok
Test Value[a]	1
Cases < Test Value	3
Cases >= Test Value	7
Total Cases	10
Number of Runs	6
Z	.245
Asymp. Sig. (2-tailed)	.806

a. Median

Run Test

- **Kesimpulan**

Berdasarkan hasil analisis diperoleh jumlah runs sebanyak 7 sedangkan tingkat signifikansi sebesar 0,806. Karena jumlah runs sebanyak 7 terletak di antara harga r kecil (3) dan harga r besar (10), atau karena nilai Asymp Sig.(0,806) lebih besar dari alpha (0,05), maka hipotesis nol tidak dapat ditolak, sehingga hipotesis yang menyatakan *"Kesukaan pemuda desa suka ribut suka merokok bersifat acak",* **diterima.**

Latihan soal 1

Diambil 30 orang mahasiswa yang mengambil mata kuliah Statistika Non-Parametrik beserta nilai UTS-nya. Akan diteliti apakah pengambilan sampel nilai UTS ini bersifat acak? Gunakan alpha = 5%. Data yang diperoleh sbb:

ID	1	2	3	4	5	6	7	8	9	10	11	12	12	14	15
Nilai Ujian	65	45	49	74	80	90	64	57	68	54	76	72	64	52	90
ID	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Nilai Ujian	94	58	60	58	69	79	83	66	62	82	84	52	41	62	76

Sign Test
Satu Sampel

Sign Test

- Digunakan untuk menguji apakah median pada suatu data sama dengan median yang diharapkan.
- Diberi nama uji tanda (*sign test*) karena data diubah menjadi serangkaian tanda (+) dan (-).
- Asumsi :
  - Skala data minimal ordinal
  - Variabel yang diamati kontinyu

Sign Test

- Data diubah menjadi serangkaian tanda (+) dan tanda (-) yang diperoleh dari : $(X_i - M_0)$, sehingga dihasilkan n tanda.
- Apabila $(X_i - M_0) = 0$ maka diabaikan dan banyaknya yang bertanda adalah n-1.
- Skala pengukuran minimal skala ordinal.
- Sampel merupakan sampel acak dari suatu populasi dengan median M tidak diketahui.
- Nilai pengamatan sampel, diberi notasi X_i , i = 1, 2, 3, ..., n.

Statistik Uji

$$X_i - M_0 = \begin{cases} 0 & \rightarrow & \text{abaikan} \\ - & \rightarrow & S- \\ + & \rightarrow & S+ \end{cases}$$

Sign Test

Perumusan Hipotesis		
Dua sisi	Satu sisi kiri	Satu sisi kanan
H_0 : $M = M_0$	H_0 : $M \geq M_0$	H_0 : $M \leq M_0$
H_1 : $M \neq M_0$	H_1 : $M < M_0$	H_1 : $M > M_0$
Statistik Uji		
$S = S' = \min \Sigma(S+ , S-)$	$S = S+ = \Sigma(S+)$	$S = S- = \Sigma(S-)$
Daerah Penolakan, Tolak H_0 jika		
$P(X \leq S' \mid b(n , 0.5)) \leq \alpha/2$	$P(X \leq S+ \mid b(n , 0.5)) \leq \alpha$	$P(X \leq S- \mid b(n , 0.5)) \leq \alpha$
$P(X \leq S \mid b(n , 05)) = \ldots$ adalah nilai peluang yang diperoleh dari tabel distribusi binomial.		

Sign Test

Aproksimasi Sampel Besar :

Untuk sampel-sampel yang berukuran **n ≥ 30** maka statistik uji:

$$\boxed{Z = \frac{(k + 0.5) - 0.5n}{0.5\sqrt{n}}}$$

Gunakan : k + 0.5, jika $k < \frac{n}{2}$

k – 0.5, jika $k > \frac{n}{2}$

P(Z < z) = ... (dari tabel normal baku)

Tolak H_0 dengan membandingkan P(Z < z) dengan taraf nyata untuk uji dua sisi dengan α/2, uji satu sisi dengan α.

Sign Test

Contoh :

Di bawah ini adalah waktu belajar mandiri dari tujuh mahasiswa. Ujilah apakah benar bahwa mahasiswa pada umumnya menyediakan waktu kurang dari dua jam untuk belajar mandiri! Gunakan taraf nyata 5%.

Mahasiswa ke-	1	2	3	4	5	6	7
Lama belajar mandiri (jam)	1.5	2.1	1.7	1.8	2.2	1.1	0.8

Sign Test

Hipotesis :

H_0 : $M \geq 2$

H_1 : $M < 2$

Statistik Uji:

- ✓ Sesuai dengan hipotesis di atas, statistik uji yang akan digunakan adalah S+, yaitu banyaknya selisih Xi – M0 yang lebih besar dari 0.
- ✓ Dari data di atas, ada dua pengamatan yang selisihnya lebih besar dari 0, yaitu mahasiswa ke-2 dan ke-5, sehingga S+ = 2.

Sign Test

Keputusan :

Dari tabel binomial, diperoleh

$P(X \leq 2 \,|\, b\,(7\,,\,0.5)) = 0.2266$

Karena $P > \alpha$ (0.05), maka gagal tolak H_0.

7	0	0.4783	0.2097	0.1335	0.0824	0.0280	0.0078	0.0016	0.0002	0.0000	
	1	0.8503	0.5767	0.4449	0.3294	0.1586	0.0625	0.0188	0.0038	0.0004	0.0000
	2	0.9743	0.8520	0.7564	0.6471	0.4199	0.2266	0.0963	0.0288	0.0047	0.0002
	3	0.9973	0.9667	0.9294	0.8740	0.7102	0.5000	0.2898	0.1260	0.0333	0.0027
	4	0.9998	0.9953	0.9871	0.9712	0.9037	0.7734	0.5801	0.3529	0.1480	0.0257
	5	1.0000	0.9996	0.9987	0.9962	0.9812	0.9375	0.8414	0.6706	0.4233	0.1497
	6		1.0000	0.9999	0.9998	0.9984	0.9922	0.9720	0.9176	0.7903	0.5217
	7			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Kesimpulan:

Tidak ada bukti kuat untuk menyatakan waktu mahasiswa untuk belajar kurang dari 2 jam.

Table A.1 Binomial Probability Sums $\sum_{x=0}^{r} b(x;n,p)$

		p									
n	r	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90
1	0	0.9000	0.8000	0.7500	0.7000	0.6000	0.5000	0.4000	0.3000	0.2000	0.1000
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	0.8100	0.6400	0.5625	0.4900	0.3600	0.2500	0.1600	0.0900	0.0400	0.0100
	1	0.9900	0.9600	0.9375	0.9100	0.8400	0.7500	0.6400	0.5100	0.3600	0.1900
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.7290	0.5120	0.4219	0.3430	0.2160	0.1250	0.0640	0.0270	0.0080	0.0010
	1	0.9720	0.8960	0.8438	0.7840	0.6480	0.5000	0.3520	0.2160	0.1040	0.0280
	2	0.9990	0.9920	0.9844	0.9730	0.9360	0.8750	0.7840	0.6570	0.4880	0.2710
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.6561	0.4096	0.3164	0.2401	0.1296	0.0625	0.0256	0.0081	0.0016	0.0001
	1	0.9477	0.8192	0.7383	0.6517	0.4752	0.3125	0.1792	0.0837	0.0272	0.0037
	2	0.9963	0.9728	0.9492	0.9163	0.8208	0.6875	0.5248	0.3483	0.1808	0.0523
	3	0.9999	0.9984	0.9961	0.9919	0.9744	0.9375	0.8704	0.7599	0.5904	0.3439
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.5905	0.3277	0.2373	0.1681	0.0778	0.0313	0.0102	0.0024	0.0003	0.0000
	1	0.9185	0.7373	0.6328	0.5282	0.3370	0.1875	0.0870	0.0308	0.0067	0.0005
	2	0.9914	0.9421	0.8965	0.8369	0.6826	0.5000	0.3174	0.1631	0.0579	0.0086
	3	0.9995	0.9933	0.9844	0.9692	0.9130	0.8125	0.6630	0.4718	0.2627	0.0815
	4	1.0000	0.9997	0.9990	0.9976	0.9898	0.9688	0.9222	0.8319	0.6723	0.4095
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.5314	0.2621	0.1780	0.1176	0.0467	0.0156	0.0041	0.0007	0.0001	0.0000
	1	0.8857	0.6554	0.5339	0.4202	0.2333	0.1094	0.0410	0.0109	0.0016	0.0001
	2	0.9842	0.9011	0.8306	0.7443	0.5443	0.3438	0.1792	0.0705	0.0170	0.0013
	3	0.9987	0.9830	0.9624	0.9295	0.8208	0.6563	0.4557	0.2557	0.0989	0.0159
	4	0.9999	0.9984	0.9954	0.9891	0.9590	0.8906	0.7667	0.5798	0.3446	0.1143
	5	1.0000	0.9999	0.9998	0.9993	0.9959	0.9844	0.9533	0.8824	0.7379	0.4686
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	0.4783	0.2097	0.1335	0.0824	0.0280	0.0078	0.0016	0.0002	0.0000	
	1	0.8503	0.5767	0.4449	0.3294	0.1586	0.0625	0.0188	0.0038	0.0004	0.0000
	2	0.9743	0.8520	0.7564	0.6471	0.4199	0.2266	0.0963	0.0288	0.0047	0.0002
	3	0.9973	0.9667	0.9294	0.8740	0.7102	0.5000	0.2898	0.1260	0.0333	0.0027
	4	0.9998	0.9953	0.9871	0.9712	0.9037	0.7734	0.5801	0.3529	0.1480	0.0257
	5	1.0000	0.9996	0.9987	0.9962	0.9812	0.9375	0.8414	0.6706	0.4233	0.1497
	6		1.0000	0.9999	0.9998	0.9984	0.9922	0.9720	0.9176	0.7903	0.5217
	7			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Latihan soal 2

As the manager of a large retail store, you have spent a lot of money on advertising to address potential customer awareness of your store. You now say that the median awareness score is 100 (whatever that means). Is your opinion correct? In support of your opinion, you randomly select 20 customers as they are exiting some other store and ask for an overall awareness score. The data are: 90 110 105 85 95 100 105 120 150 85 100 110 90 75 105 120 100 105 90 100

(a) State the null and alternative hypotheses.

(b) Is your opinion correct or not? Show your analysis and how you reached your conclusion.

Sign Test
2 Sampel

Sign Test 2 Sampel

Merupakan perluasan dari uji tanda untuk satu sampel

Asumsi :

- ✓ Data pengamatan berupa sampel acak dari n pasangan hasil pengukuran : (X_i, Y_i); $i = 1,2,....,n$
- ✓ Hasil pengukuran dapat diperoleh dari satu subyek atau subyek yang berbeda yang telah dipasangkan
- ✓ Hasil pengukuran masing-masing pasangan adalah saling bebas

Sign Test 2 Sampel

- **Hipotesis Dua Arah**

  H_0 : Median populasi beda (selisih),(Xi,Yi) = Di = 0 atau

  H_0 : $P(+)=P(-)=0.5$

  H_1 : Median populasi beda (selisih),(Xi,Yi) = Di ≠ 0 atau

  H_1 : $P(+)\neq P(-)=0.5$

- **Statistik uji**: $S = S' = \min(\sum(+), \sum(-))$
- **Daerah Penolakan**: Tolak H_0 jika $P(X \leq S' \mid b(n, 0.5)) \leq \alpha/2$

Sign Test 2 Sampel

- **Hipotesis Satu Arah Kanan**

  H_0 : Median populasi beda (selisih),(Xi,Yi) = Di ≤ 0 atau

  H_0 : P(+) ≤ P(-)

  H_1 : Median populasi beda (selisih),(Xi,Yi) = Di > 0 atau

  H_1 : P(+) > P(-)

- **Statistik uji**: S = S- = ∑ (-)
- **Daerah Penolakan**: Tolak H_0 jika $P(X \leq S- \mid b(n , 0.5)) \leq \alpha$

Sign Test 2 Sampel

- **Hipotesis Satu Arah Kiri**

H_0 : Median populasi beda (selisih),(Xi,Yi) = Di ≥ 0 atau

H_0 : $P(+) \geq P(-)$

H_1 : Median populasi beda (selisih),(Xi,Yi) = Di < 0 atau

H_1 : $P(+) < P(-)$

- **Statistik uji**: $S = S+ = \sum (+)$
- **Daerah Penolakan**: Tolak H_0 jika $P(X \leq S+ \mid b(n\ ,\ 0.5)) \leq \alpha$

Sign Test 2 Sampel

- **Aproksimasi Sampel Besar** Lihat Uji Untuk 1 Sampel
- Sifat Lainnya Sama Dengan Uji Tanda Untuk 1 Sampel

Sign Test 2 Sampel

Contoh

Latane dan Cappel mempelajari efek kebersamaan dari denyut jantung pada tikus. Mereka mencatat denyut jantung 10 ekor tikus, baik ketika masing-masing tikus itu sedang sendiri maupun ketika sedang bersama-sama. Hasil-hasil studi tersebut tedapat dalam Tabel 1. Apakah kita dapat menyimpulkan bahwa kebersamaan meningkatkan denyut jantung tikus dengan taraf nyata 0,05.

Tabel 1

Tikus	1	2	3	4	5	6	7	8	9	10
Kerika sendiri (X)	463	462	462	456	450	426	418	415	409	402
Ketika bersama (Y)	523	494	461	535	476	454	448	408	470	437

Jawaban

Hipotesis

$H_0 : P(+) \leq P(-)$

$H_1 : \; P(+) > P(-)$

Statistik uji

i	1	2	3	4	5	6	7	8	9	10
$Y_i - X_i$	60	32	-1	79	26	28	30	-7	61	35
tanda	+	+	-	+	+	+	+	-	+	+

$P(X \leq 2|10, 0{,}5) = 0{,}0547$ (buku tabel binomial)

Keputusan

Nilai 0,0547 lebih dari 0,05 sehingga tidak cukup bukti menolak H0 pada taraf nyata 0,05.

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	***r***	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
8	**0**	0.4305	0.1678	0.1001	0.0576	0.0168	0.0039	0.0007	0.0001	0.0000	
	1	0.8131	0.5033	0.3671	0.2553	0.1064	0.0352	0.0085	0.0013	0.0001	
	2	0.9619	0.7969	0.6785	0.5518	0.3154	0.1445	0.0498	0.0113	0.0012	0.0000
	3	0.9950	0.9437	0.8862	0.8059	0.5941	0.3633	0.1737	0.0580	0.0104	0.0004
	4	0.9996	0.9896	0.9727	0.9420	0.8263	0.6367	0.4059	0.1941	0.0563	0.0050
	5	1.0000	0.9988	0.9958	0.9887	0.9502	0.8555	0.6846	0.4482	0.2031	0.0381
	6		0.9999	0.9996	0.9987	0.9915	0.9648	0.8936	0.7447	0.4967	0.1869
	7		1.0000	1.0000	0.9999	0.9993	0.9961	0.9832	0.9424	0.8322	0.5695
	8				1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	**0**	0.3874	0.1342	0.0751	0.0404	0.0101	0.0020	0.0003	0.0000		
	1	0.7748	0.4362	0.3003	0.1960	0.0705	0.0195	0.0038	0.0004	0.0000	
	2	0.9470	0.7382	0.6007	0.4628	0.2318	0.0898	0.0250	0.0043	0.0003	0.0000
	3	0.9917	0.9144	0.8343	0.7297	0.4826	0.2539	0.0994	0.0253	0.0031	0.0001
	4	0.9991	0.9804	0.9511	0.9012	0.7334	0.5000	0.2666	0.0988	0.0196	0.0009
	5	0.9999	0.9969	0.9900	0.9747	0.9006	0.7461	0.5174	0.2703	0.0856	0.0083
	6	1.0000	0.9997	0.9987	0.9957	0.9750	0.9102	0.7682	0.5372	0.2618	0.0530
	7		1.0000	0.9999	0.9996	0.9962	0.9805	0.9295	0.8040	0.5638	0.2252
	8			1.0000	1.0000	0.9997	0.9980	0.9899	0.9596	0.8658	0.6126
	9					1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	**0**	0.3487	0.1074	0.0563	0.0282	0.0060	0.0010	0.0001	0.0000		
	1	0.7361	0.3758	0.2440	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	
	2	0.9298	0.6778	0.5256	0.3828	0.1673	0.0547	0.0123	0.0016	0.0001	
	3	0.9872	0.8791	0.7759	0.6496	0.3823	0.1719	0.0548	0.0106	0.0009	0.0000
	4	0.9984	0.9672	0.9219	0.8497	0.6331	0.3770	0.1662	0.0473	0.0064	0.0001
	5	0.9999	0.9936	0.9803	0.9527	0.8338	0.6230	0.3669	0.1503	0.0328	0.0016
	6	1.0000	0.9991	0.9965	0.9894	0.9452	0.8281	0.6177	0.3504	0.1209	0.0128
	7		0.9999	0.9996	0.9984	0.9877	0.9453	0.8327	0.6172	0.3222	0.0702
	8		1.0000	1.0000	0.9999	0.9983	0.9893	0.9536	0.8507	0.6242	0.2639
	9				1.0000	0.9999	0.9990	0.9940	0.9718	0.8926	0.6513
	10					1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

LATIHAN

Shani dkk mempelajari efek fenobarbital terhadap fungsi hati pada pasien-pasien penderita sindrom Dubin-Johnson. Tabel berikut menunjukkan kadar bilirubin dalam serum pasien-pasien pada sebelum dan sesudah pemberian fenobarbital. Berdasarkan data ini, dapatkah kita menyimpulkan bahwa fenobarbital mengurangi kadar bilirubin? Misalkan tingkat signifikansi 0,05.

Pasien	1	2	3	4	5	6	7	8	9	10	11	12	13
Sebelum (X)	4,0	3,2	3,8	1,8	3,0	5,3	5,7	3,0	2,7	2,9	2,8	1,8	2,6
Sesusah (Y)	3,1	3,0	3,5	1,0	1,8	3,9	2,2	2,1	1,4	2,9	2,6	1,4	2,5

Terima kasih

Statistika Non Parametrik
TSD - Ganjil 2023/2024

Minggu Ke-3:
“Binomial Test dan Wilcoxon Test”

Pedoman Memilih Teknik Statistik Non Parametrik Pengujian Hipotesis Satu Sampel

Skala yang Digunakan	Alat Analisis Pengujian Hipotesis Satu Sampel
Nominal	Uji Run Uji Sign Uji Binomial
Ordinal	Uji Wilcoxon Signed Rank

Binomial Test

Binomial Test

Ketentuan

1. Binomial test digunakan untuk menguji satu populasi yang mempunyai hanya dua grup.
2. Skala data untuk populasinya adalah skala Nominal
3. Proporsi grup pertama adalah p dan proporsi grup kedua adalah $q = 1 - p$.

Binomial Test

Fungsi:

Menguji perbedaan proporsi untuk grup-grup pada populasi, yang mana sampel yang digunakan adalah sampel tunggal.

Bila *n* adalah banyaknya seluruh observasi, maka peluang untuk mendapatkan x observasi dalam salah satu grup dan mendapatkan (n − x) observasi dalam grup lainnya adalah

$$P(x) = \binom{n}{x} p^x (q)^{n-x}$$

dengan:

p =proporsi yang diharapkan untuk salah satu grup

q = 1 − p

Binomial Test

Tahap Pengujian

1. Menentukan hipotesis

   $H_0 : p = \frac{1}{2}$

   $H_1 : p \neq \frac{1}{2};\ H_1 : p < \frac{1}{2};\ H_1 : p > \frac{1}{2}$

2. Menentukan tingkat signifikansi α
3. Menentukan uji statistik
   a. Menentukan banyaknya observasi N
   b. Menentukan banyaknya observasi pada grup pertama dan pada grup kedua, masing-masing diberikan notasi n_p dan n_q
   c. Jika banyaknya observasi ≤ 25 dan $p = q = \frac{1}{2}$ maka dapat digunakan **Tabel Binomial**

   p-value$= \sum_{i=1}^{x} \binom{N}{i} p^i q^{N-i}$

   $x = \min(n_p, n_q)$

   d. Jika banyaknya observasi ≤ 25 dan $p \neq \frac{1}{2}$, maka dapat menggunakan **Table of Binomial** tapi terbatas untuk $N = 20$
   e. Jika observasi berukuran > 25 dapat didekati dengan distribusi Normal

$$z = \frac{(x \pm 0{,}5) - Np}{\sqrt{Npq}} \begin{cases} x + 0{,}5, jika\ x < NP \\ x - 0{,}5, jika\ x \geq NP \end{cases}$$

   Nilai z yang diperoleh digunakan untuk menghitung nilai p-value

4. Menentukan daerah penolakan

   Tolak H_0 jika p-value$< \alpha$

5. Mengambil kesimpulan

Binomial Test

Contoh 1

Dilakukan penelitian untuk mengetahui kecenderungan masyarakat dalam memilih perawatan kecantikan. Berdasarkan 20 anggota sampel yang dipilih secara acak, ternyata 8 orang memilih perawatan kecantikan di salon dan 12 lainnya lebih memilih klinik kecantikan.

Ujilah bahwa peluang masyarakat dalam memilih perawatan kecantikan di salon dan di klinik kecantikan adalah sama! Gunakan taraf signifikansi 5%.

Solusi

Misalkan A: salon dan B: klinik kecantikan

1. $H_0 : p = q = \frac{1}{2}$

   $H_1 : p \neq q$
2. $\alpha = 5\%$
3. $N = 20, n_A = 8, n_B = 12$

   $x = \min(n_A, n_B) = \min(8,12) = 8$

   p-value$= \sum_{i=1}^{8} \binom{N}{i} p^i q^{N-i} = \boxed{0{,}252}$

   Karena dua sisi maka $2 \times p - value = 2 \times 0{,}252 = 0{,}504$
4. Gagal tolak H_0 karena p-value> α
5. Peluang masyarakat memilih salon dan klinik kecantikan adalah sama

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	r	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
19	**0**	0.1351	0.0144	0.0042	0.0011	0.0001					
	1	0.4203	0.0829	0.0310	0.0104	0.0008	0.0000				
	2	0.7054	0.2369	0.1113	0.0462	0.0055	0.0004	0.0000			
	3	0.8850	0.4551	0.2631	0.1332	0.0230	0.0022	0.0001			
	4	0.9648	0.6733	0.4654	0.2822	0.0696	0.0096	0.0006	0.0000		
	5	0.9914	0.8369	0.6678	0.4739	0.1629	0.0318	0.0031	0.0001		
	6	0.9983	0.9324	0.8251	0.6655	0.3081	0.0835	0.0116	0.0006		
	7	0.9997	0.9767	0.9225	0.8180	0.4878	0.1796	0.0352	0.0028	0.0000	
	8	1.0000	0.9933	0.9713	0.9161	0.6675	0.3238	0.0885	0.0105	0.0003	
	9		0.9984	0.9911	0.9674	0.8139	0.5000	0.1861	0.0326	0.0016	
	10		0.9997	0.9977	0.9895	0.9115	0.6762	0.3325	0.0839	0.0067	0.0000
	11		1.0000	0.9995	0.9972	0.9648	0.8204	0.5122	0.1820	0.0233	0.0003
	12			0.9999	0.9994	0.9884	0.9165	0.6919	0.3345	0.0676	0.0017
	13			1.0000	0.9999	0.9969	0.9682	0.8371	0.5261	0.1631	0.0086
	14				1.0000	0.9994	0.9904	0.9304	0.7178	0.3267	0.0352
	15					0.9999	0.9978	0.9770	0.8668	0.5449	0.1150
	16					1.0000	0.9996	0.9945	0.9538	0.7631	0.2946
	17						1.0000	0.9992	0.9896	0.9171	0.5797
	18							0.9999	0.9989	0.9856	0.8649
	19							1.0000	1.0000	1.0000	1.0000
20	**0**	0.1216	0.0115	0.0032	0.0008	0.0000					
	1	0.3917	0.0692	0.0243	0.0076	0.0005	0.0000				
	2	0.6769	0.2061	0.0913	0.0355	0.0036	0.0002				
	3	0.8670	0.4114	0.2252	0.1071	0.0160	0.0013	0.0000			
	4	0.9568	0.6296	0.4148	0.2375	0.0510	0.0059	0.0003			
	5	0.9887	0.8042	0.6172	0.4164	0.1256	0.0207	0.0016	0.0000		
	6	0.9976	0.9133	0.7858	0.6080	0.2500	0.0577	0.0065	0.0003		
	7	0.9996	0.9679	0.8982	0.7723	0.4159	0.1316	0.0210	0.0013	0.0000	
	8	0.9999	0.9900	0.9591	0.8867	0.5956	0.2517	0.0565	0.0051	0.0001	
	9	1.0000	0.9974	0.9861	0.9520	0.7553	0.4119	0.1275	0.0171	0.0006	
	10		0.9994	0.9961	0.9829	0.8725	0.5881	0.2447	0.0480	0.0026	0.0000
	11		0.9999	0.9991	0.9949	0.9435	0.7483	0.4044	0.1133	0.0100	0.0001
	12		1.0000	0.9998	0.9987	0.9790	0.8684	0.5841	0.2277	0.0321	0.0004
	13			1.0000	0.9997	0.9935	0.9423	0.7500	0.3920	0.0867	0.0024
	14				1.0000	0.9984	0.9793	0.8744	0.5836	0.1958	0.0113
	15					0.9997	0.9941	0.9490	0.7625	0.3704	0.0432
	16					1.0000	0.9987	0.9840	0.8929	0.5886	0.1330
	17						0.9998	0.9964	0.9645	0.7939	0.3231
	18						1.0000	0.9995	0.9924	0.9308	0.6083
	19							1.0000	0.9992	0.9885	0.8784
	20								1.0000	1.0000	1.0000

Wilcoxon Signed Rank Test for 1 Population

Wilcoxon Signed Rank Test

- Uji wilcoxon merupakan uji peringkat bertanda.
- Pada uji ini dilakukan pada hasil selisih data dengan median yang dihipotesiskan (tanpa memperhatikan tanda).
- Setelah diperingkat baru diberi tanda sesuai hasil selisihnya.

Wilcoxon Signed Rank Test

Ketentuan

1. Null Hypothesis

   $$H_0 : \tilde{\mu} = \tilde{\mu}_0$$

2. Hitung **selisih** masing-masing sampel terhadap nilai $\tilde{\mu}_0$. Selisih yang nilainya **nol** tidak diikutkan dalam perhitungan. Dan $\boldsymbol{n}$ adalah banyaknya data selisih

3. Ranking 1 diberikan untuk nilai **mutlak selisih yang paling kecil**, ranking 2 diberikan untuk nilai **mutlak selisih paling kecil kedua**, dan seterusnya.

4. Ketika ditemukan nilai mutlak yang **sama**, maka rankingnya diberikan nilai **rata-rata** dari ranking keduanya. Contoh: ranking ke-5 dan ranking ke-6 memiliki nilai nilai mutlak selisih yang sama sehingga ranking untuk keduanya adalah 5,5.

5. Hitung total ranking untuk selisih yang bernilai negatif dan berikan notasi $\boldsymbol{w}_-$ dan hitung total ranking untuk selisih yang bernilai positif dan berikan notasi $\boldsymbol{w}_+$, kemudian $\boldsymbol{min}(\boldsymbol{w}_-, \boldsymbol{w}_+) = \boldsymbol{w}$

H_0	H_1		Compute
$\tilde{\mu} = \tilde{\mu}_0$	$\tilde{\mu} < \tilde{\mu}_0$	One side	w_+
	$\tilde{\mu} > \tilde{\mu}_0$	One side	w_-
	$\tilde{\mu} \neq \tilde{\mu}_0$	Two side	w
$\tilde{\mu}_1 = \tilde{\mu}_2$	$\tilde{\mu}_1 < \tilde{\mu}_2$	One side	w_+
	$\tilde{\mu}_1 > \tilde{\mu}_2$	One side	w_-
	$\tilde{\mu}_1 \neq \tilde{\mu}_2$	Two side	w

Tolak $\boldsymbol{H_0}$, jika $\boldsymbol{w}_-$ $\boldsymbol{atau}$ $\boldsymbol{w}_+$ $\boldsymbol{atau}$ $\boldsymbol{w} \leq$ titik kritis

Wilcoxon Signed Rank Test

Data berikut ini merupakan jumlah jam dari suatu pengisian tabung oksigen:

1,5; 2,2; 0,9; 1,3; 2,0; 1,6; 1,8; 1,5; 2,0; 1,2; 1,7

Dengan menggunakan Wilcoxon sign ranked test dan taraf signifikansi 5%, uji apakah jumlah jam pengisian tabung oksigen memiliki median 1,8

Solusi

1. $H_0 : \tilde{\mu} = 1{,}8$
2. $H_1 : \tilde{\mu} \neq 1{,}8$
3. $\alpha = 5\%$
4. **Perhitungan:**

sampel	1.5	2.2	0.9	1.3	2	1.6	1.8	1.5	2	1.2	1.7
selisih	-0.3	0.4	-0.9	-0.5	0.2	-0.2	0	-0.3	0.2	-0.6	-0.1
ranking	5.5	7	10	8	3	3		5.5	3	9	1
w-	42										
w+	13										
w	13										

$n = 10$, sehingga daerah kritisnya adalah $w \leq \boxed{8}$

$w = 13$ tidak berada pada daerah kritis

6. Kesimpulan: Gagal Tolak H_0, artinya median dari pengisian tabung oksigen tidak berbeda secara signifikan dari 1,8 jam

Table A.16 Critical Values for the Signed-Rank Test

n	One-Sided $\alpha = 0.01$ Two-Sided $\alpha = 0.02$	One-Sided $\alpha = 0.025$ Two-Sided $\alpha = 0.05$	One-Sided $\alpha = 0.05$ Two-Sided $\alpha = 0.1$
5			1
6		1	2
7	0	2	4
8	2	4	6
9	3	6	8
10	5	$\boxed{8}$	11
11	7	11	14
12	10	14	17
13	13	17	21
14	16	21	26
15	20	25	30
16	24	30	36
17	28	35	41
18	33	40	47
19	38	46	54
20	43	52	60
21	49	59	68
22	56	66	75
23	62	73	83
24	69	81	92
25	77	90	101
26	85	98	110
27	93	107	120
28	102	117	130
29	111	127	141
30	120	137	152

Reproduced from F. Wilcoxon and R. A. Wilcox, *Some Rapid Approximate Statistical Procedures, American Cyanamid Company*, Pearl River, N.Y., 1964, by permission of the American Cyanamid Company.

Wilcoxon Signed Rank Test for Large Samples

Ketika $n \geq 15$, maka $W_+ (atau\ W_-)$ dapat didekati dengan distribusi Normal dengan mean dan variansi masing-masing adalah

$$\mu_{W_+} = \frac{n(n+1)}{4} \qquad \sigma^2_{W_+} = \frac{n(n+1)(2n+1)}{24}$$

Ketika nilai n lebih besar tidak tersedia di dalam Tabel, maka statistik berikut

$$Z = \frac{W_+ - \mu_{W_+}}{\sigma_{W_+}}$$

dapat digunakan sebagai titik kritis untuk pengujian

Wilcoxon Signed Rank Test

Latihan

Seorang dokter mengatakan bahwa ia dalam setahun melakukan visit per pasien di RS dengan nilai median 5 kali. Untuk membuktikan validitas pernyataanya, ia secara acak memilih sepuluh pasien dan menghitung jumlah visit di RS selama setahun terakhir. Data yang diperoleh untuk jumlah kunjungan per pasien selama setahun adalah 9, 10, 8, 4, 8, 3, 0, 10, 15, 9. Apakah dengan data tersebut dapat dibuktikan pernyataannya bahwa ia jumlah kunjungan per pasien oleh dokter adalah 5, dengan $\alpha = 0,05$?

Wilcoxon Signed Rank Test

Hipotesis

H_0 : $M = 5$

H_1 : $M \neq 5$

Penghitungan

Subyek	X	D = X – M	Rank of |D|	Signed rank of |D|
1	9	4	5.5	5.5
2	10	5	8	8
3	8	3	3.5	3.5
4	4	-1	1	-1
5	8	3	3.5	3.5
6	3	-2	2	-2
7	0	-5	8	-8
8	10	5	8	8
9	15	10	10	10
10	9	4	5.5	5.5
			∑ T+	**44**
			∑ T-	**11**

Wilcoxon Signed Rank Test

Keputusan

Dengan nilai $\alpha = 0.05$ dan $n = 10$ dan menggunakan tabel nilai T kritis wilcoxon signed-ranks maka diperoleh nilai Ttabel uji hipotesa dua arah (two tailed) = 8, maka:

Untuk uji hipotesis dua arah (two tailed) Ttabel = 8 < Thitung = 11, sehingga hipotesis nol tidak ditolak atau jumlah kunjungan per pasien oleh dokter 5 kali.

Kesimpulan

Dapat disimpulkan data yang dikumpulkan mengindikasikan bahwa sampel 10 obyek berasal dari populasi dengan nilai median 5 atau cukup bukti bahwa dokter melakukan kunjungan per pasien dengan median = 5.

Wilcoxon Signed Rank Test

Seorang dosen beranggapan bahwa median IP mahasiswa suatu kelas pada semester tertentu kurang dari 3.40. Ujilah anggapan dosen tersebut jika IP dari 10 orang mahasiswa yang diambil secara acak dari kelas tersebut adalah seperti yang tersaji dalam tabel berikut : (Gunakan taraf nyata 5%)

Mahasiswa ke	1	2	3	4	5	6	7	8	9	10
IP	3.35	3.45	3.30	3.25	3.52	3.38	3.10	3.42	3.42	3.38

Wilcoxon Test for 2 Population

Wilcoxon Test

Merupakan perluasan dari uji wilcoxon untuk 1 sampel.

Asumsi :

- Data pengamatan berupa sampel acak dari n pasangan (X_i, Y_i) ; i = 1,2,....,n yang diperoleh dari hasil pengukuran subyek yang sama atau subyek berbeda yang telah dipasangkan
- Hasil pengukuran tiap pasangan saling bebas
- Skala pengukuran minimal ordinal

Wilcoxon Test

- **Hipotesis Dua Arah**

  H_0 : Median populasi beda (selisih) = 0 atau $M_D = 0$

  H_1 : Median populasi beda (selisih) ≠ 0 atau $M_D \neq 0$

- **Statistik uji**: T = T' = min (T+ , T-)
- **Daerah Penolakan**: Tolak H_0 jika $T' < T\ n,\alpha/2$

Wilcoxon Test

- **Hipotesis Satu Arah Kanan**

H_0 : Median populasi beda (selisih) ≤ 0 atau $M_D \leq 0$

H_1 : Median populasi beda (selisih) > 0 atau $M_D > 0$

- **Statistik uji**: T = T-
- **Daerah Penolakan**: tolak H_0 jika T- < T n,α/2

Wilcoxon Test

- **Hipotesis Satu Arah Kiri**

H_0 : Median populasi beda (selisih) ≥ 0 atau $M_D \geq 0$

H_1 : Median populasi beda (selisih) < 0 atau $M_D < 0$

- **Statistik uji**: T = T+
- **Daerah Penolakan**: tolak H_0 jika $T+ < T\ n,\alpha/2$

Wilcoxon Test

Contoh:

Dickie dkk mengkaji perubahan-perubahan hermodinamik pada pasien-pasien dengan *pulmonary thromboembolism* yang akut. Tabel berikut memperlihatkan tekanan arteri paru rata-rata yang teramati pada sembilan orang dari pasien-pasien sebelum dan 24 jam setelah terapi urokinase. Kita ingin tahu apakah data ini menyediakan bukti yang cukup untuk menunjukkan bahwa terapi urokinase menurunkan tekanan arteri paru? (Misalkan $\alpha = 0{,}05$)

Pasien	1	2	3	4	5	6	7	8	9
0 jam (X)	33	17	30	25	36	25	31	20	18
24 jam (Y)	21	17	22	13	33	20	19	13	9

Jawaban

Hipotesis

H_0 : Median populasi beda (selisih) ≥ 0 atau $M_D \geq 0$

H_1 : Median populasi beda (selisih) < 0 atau $M_D < 0$

X	Y	$D = X - Y$	$\|D\|$
33	21	12	7
17	17	0	abaikan
30	22	8	4
25	13	12	7
36	33	3	1
25	20	5	2
31	19	12	7
20	13	7	3
18	9	9	5
			T+ = 36

Dari Tabel uji peringkat bertanda Wilcoxon dengan n=8 menunjukkan bahwa peluang untuk mendapatkan suatu nilai T+ = 36, adalah **≈>0,5273**, dimana Nilai tersebut lebih besar dari 0,05 sehingga kita cukup bukti untuk **gagal menolak H_0**.

Jadi, kita dapat menyimpulkan bahwa median populasi beda (selisih) sama dengan nol, artinya terapi urokinase dapat menurunkan tekanan arteri paru.

Catatan: Nilai d didapatkan dari Tabel A.3 di slide berikutnya

Jawaban

```
x=c(33,17,30,25,36,25,31,20,18)
y=c(21,17,22,13,33,20,19,13,9)
wilcox.test(x, y, paired = TRUE, alternative = "less")
```

```
        Wilcoxon signed rank test with continuity correction

data:  x and y
V = 36, p-value = 0.9954
alternative hypothesis: true location shift is less than 0
```

Kesimpulan : gagal menolak H_0
Terapi urokinase dapat menurunkan tekanan arteri paru

TABLE A.3 Probability levels for the Wilcoxon signed-rank test

T	P
***n* = 5**	
*0	.0313
1	.0625
2	.0938
3	.1563
4	.2188
5	.3125
6	.4063
7	.5000
***n* = 6**	
0	.0156
1	.0313
*2	.0469
3	.0781
4	.1094
5	.1563
6	.2188
7	.2813
8	.3438
9	.4219
10	.5000
***n* = 7**	
0	.0078
1	.0156
2	.0234
*3	.0391
4	.0547
5	.0781
6	.1094
7	.1484
8	.1875
9	.2344
10	.2891
11	.3438
12	.4063
13	.4688
14	.5313

T	P
***n* = 8**	
0	.0039
1	.0078
2	.0117
3	.0195
4	.0273
*5	.0391
6	.0547
7	.0742
8	.0977
9	.1250
10	.1563
11	.1914
12	.2305
13	.2734
14	.3203
15	.3711
16	.4219
17	.4727
18	.5273
***n* = 9**	
0	.0020
1	.0039
2	.0059
3	.0098
4	.0137
5	.0195
6	.0273
7	.0371
*8	.0488
9	.0645
10	.0820
11	.1016
12	.1250
13	.1504
14	.1797
15	.2129
16	.2480
17	.2852
18	.3262
19	.3672
20	.4102
21	.4551
22	.5000

T	P
***n* = 10**	
0	.0010
1	.0020
2	.0029
3	.0049
4	.0068
5	.0098
6	.0137
7	.0186
8	.0244
9	.0322
*10	.0420
11	.0527
12	.0654
13	.0801
14	.0967
15	.1162
16	.1377
17	.1611
18	.1875
19	.2158
20	.2461
21	.2783
22	.3125
23	.3477
24	.3848
25	.4229
26	.4609
27	.5000

T	P
***n* = 11**	
0	.0005
1	.0010
2	.0015
3	.0024
4	.0034
5	.0049
6	.0068
7	.0093
8	.0122
9	.0161
10	.0210
11	.0269
12	.0337
*13	.0415
14	.0508
15	.0615
16	.0737
17	.0874
18	.1030
19	.1201
20	.1392
21	.1602
22	.1826
23	.2065
24	.2324
25	.2598
26	.2886
27	.3188
28	.3501
29	.3823
30	.4155
31	.4492
32	.4829
33	.5171

T	P
***n* = 12**	
0	.0002
1	.0005
2	.0007
3	.0012
4	.0017
5	.0024
6	.0034
7	.0046
8	.0061
9	.0081
10	.0105
11	.0134
12	.0171
13	.0212
14	.0261
15	.0320
16	.0386
*17	.0461
18	.0549
19	.0647
20	.0757
21	.0881
22	.1018
23	.1167
24	.1331
25	.1506
26	.1697
27	.1902
28	.2119
29	.2349
30	.2593
31	.2847
32	.3110
33	.3386
34	.3667
35	.3955
36	.4250
37	.4548
38	.4849
39	.5151

T	P
***n* = 13**	
0	.0001
1	.0002
2	.0004
3	.0006
4	.0009
5	.0012
6	.0017
7	.0023
8	.0031
9	.0040
10	.0052
11	.0067
12	.0085
13	.0107
14	.0133
15	.0164
16	.0199
17	.0239
18	.0287
19	.0341
20	.0402
*21	.0471
22	.0549
23	.0636
24	.0732
25	.0639
26	.0955
27	.1082
28	.1219
29	.1367
30	.1527
31	.1698
32	.1879
33	.2072
34	.2274
35	.2487
36	.2709
37	.2939
38	.3177
39	.3424
40	.3677
41	.3934
42	.4197
43	.4463
44	.4730
45	.5000


Terima kasih

Statistika Non Parametrik
TSD - Ganjil 2023/2024

Minggu Ke-4:
"Chi-square Test"

Chi-square Test

Chi-square Test

- Homogenitas
- Ketidaktergantungan / independendensi

Chi-Square Test

Asumsi:

- Data terdiri atas sebuah sampel acak sederhana dengan ukuran *n* dari suatu populasi yang diamati
- Hasil-hasil pengamatan dalam sampel dapat diklasifikasikan secara silang menjadi dua kriteria, sehingga masing-masing hasil pengamatan hanya memiliki satu kriteria saja dari masing-masing kriteria. Kriteria disini meupakan variabel-variabel yang diamati

Chi-Square Test

Kkriteria Klasifikasi Pertama	Kriteria Klasifikasi Kedua Tingkat						
Tingkat	1	2	…	j	…	c	Jumlah
1	O11	O12		O1j		O1c	n1.
2	O21	O22		O2j		O2c	n2.
. . .							
j	Oj1	Oj2		Ojj		Ojc	nj.
. . .							
r	Or1	Or2		Orj		Orj	nr.
Jumlah	n.1	n.2		n.j		n.c	n

Chi-Square Test

Hipotesis

H0 : Kedua kriteria klasifikasi saling bebas atau independent

H1 : Kedua kriteria klasifikasi tidak saling bebas

Statistik Uji: $$\chi^2 = \sum_{\text{cells}} \frac{(O - E)^2}{E}$$

Dengan

O: observed cell frequency

E: expected cell frequency = $\frac{row\ total\ x\ column\ total}{Grand\ total}$

Degree of freedom (df): $(r-1)(c-1)$ dimana r adalah jumlah baris dan c adalah jumlah kolom tabel kontingensi

Kriteria Penolakan

- Jika nilai statistik uji (χ^2) lebih besar dari nilai tabel (${\chi^2}_{\alpha}$), maka tolak H0.
- Jika nilai statistik uji lebih kecil dari nilai tabel, maka gagal tolak H0

Chi-Square Test

Hipotesis

H0 : Populasi-populasi asal sampel homogen

H1 : Populasi-populasi asal sampel tidak homogen

Statistik Uji:
$$\chi^2 = \sum_{\text{cells}} \frac{(O - E)^2}{E}$$

Dengan

O: observed cell frequency

E: expected cell frequency = $\frac{row\ total\ x\ column\ total}{Grand\ total}$

Degree of freedom (df): $(r-1)(c-1)$ dimana r adalah jumlah baris dan c adalah jumlah kolom tabel kontingensi

Kriteria Penolakan

- Jika nilai statistik uji (χ^2) lebih besar dari nilai tabel (${\chi^2}_{\alpha}$), maka tolak H0.
- Jika nilai statistik uji lebih kecil dari nilai tabel, maka gagal tolak H0

Chi-Square Test

Contoh

Pada suatu percobaan perkecambahan digunakan 250 biji. Dimana 100 biji diantaranya diberi pupuk kimia, sedangkan 150 biji lainnya tidak diberi pupuk kimia. Jumlah benih yang berkecambah dicatat pada Tabel di bawah ini. Apakah data tersebut memberikan cukup bukti untuk menyimpulkan bahwa data bersifat homogen?

Tabel Kontingensi	Berkecambah	Tidak berkecambah	Total
Diperi perlakuan	84	16	100
Tidak diberi perlakuan	132	18	150
Total	216	34	250

Chi-Square Test

1. Menghitung expected cell frequency = $\frac{row\ total\ x\ column\ total}{Grand\ total}$

Expected Value	Berkecambah	Tidak berkecambah
Diperi perlakuan	86,4	13,6
Tidak diberi perlakuan	129,6	20,4

2. Menghitung Chi-square $\chi^2 = \sum_{cells} \frac{(O - E)^2}{E}$

Chi-square	Berkecambah	Tidak berkecambah	
Diperi perlakuan	0,067	0,424	
Tidak diberi perlakuan	0,044	0,282	
		Chi hitung	0,817
		Chi-tabel	3,84

	Probability less than the critical value				
ν	0.90	0.95	0.975	0.99	0.999
1	2.706	3.841	5.024	6.635	10.828
2	4.605	5.991	7.378	9.210	13.816
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.833	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322
8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588
11	17.275	19.675	21.920	24.725	31.264
12	18.549	21.026	23.337	26.217	32.910
13	19.812	22.362	24.736	27.688	34.528
14	21.064	23.685	26.119	29.141	36.123
15	22.307	24.996	27.488	30.578	37.697
16	23.542	26.296	28.845	32.000	39.252
17	24.769	27.587	30.191	33.409	40.790
18	25.989	28.869	31.526	34.805	42.312
19	27.204	30.144	32.852	36.191	43.820
20	28.412	31.410	34.170	37.566	45.315
21	29.615	32.671	35.479	38.932	46.797
22	30.813	33.924	36.781	40.289	48.268
23	32.007	35.172	38.076	41.638	49.728
24	33.196	36.415	39.364	42.980	51.179
25	34.382	37.652	40.646	44.314	52.620
26	35.563	38.885	41.923	45.642	54.052
27	36.741	40.113	43.195	46.963	55.476
28	37.916	41.337	44.461	48.278	56.892
29	39.087	42.557	45.722	49.588	58.301
30	40.256	43.773	46.979	50.892	59.703
31	41.422	44.985	48.232	52.191	61.098
32	42.585	46.194	49.480	53.486	62.487
33	43.745	47.400	50.725	54.776	63.870
34	44.903	48.602	51.966	56.061	65.247
35	46.059	49.802	53.203	57.342	66.619

Chi-Square Test Using R

```
table1=matrix(c(84,132,16,18),ncol=2)
colnames(table1)=c("Berkecambah","Tidak Berkecambah")
rownames(table1)=c("Perlakuan","Tanpa Perlakuan")
table1

Uji=chisq.test(table1)
Uji
```

Bartlett’s Test

Bartlett's Test

- Bartlett's test dikenalkan oleh Snedecor dan Cochran pada tahun 1983
- Bartlett's test digunakan untuk melakukan pengujian *k* sampel yang memiliki varians sama. Jika sampel memiliki varians yang sama maka dapat dikatakan sampel tersebut bersifat homogen
- Kelemahan dari Bartlett's test adalah sensitive terhadap distribusi normal, sehingga alternatif pengujian lain yang dapat digunakan adalah Levene test yang cenderung tidak terlalu sensitive terhadap normality.

Bartlett's Test

Hipotesis

H0 : $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

H1 : minimal ada satu $\sigma_i^2 \neq \sigma_j^2$

Statistik Uji

$$T = \frac{(N-k)\ln s_p^2 - \sum_{i=1}^{k}(N_i - 1)\ln s_i^2}{1 + \left(1/\left(3(k-1)\right)\right)\left(\left(\sum_{i=1}^{k} 1/(N_i - 1)\right) - 1/(N-k)\right)}$$

$$s_p^2 = \sum_{i=1}^{k}(N_i - 1)s_i^2 / (N-k)$$

Dimana,
s_i^2: varians dari group ke-*i*
s_p^2: pooled variance
N : total sampel
k : jumlah group

Kriteria Penolakan

- Jika nilai statistik uji (T) lebih besar dari nilai tabel ($\chi^2{}_{\alpha,k-1}$), maka tolak H0.
- Jika nilai statistik uji lebih kecil dari nilai tabel, maka gagal tolak H0

Bartlett’s Test Using R

```
#Setting Directory
setwd("D:/UNAIR/1. Perkuliahan/Statnonpar/2023-2024/M4")

#Import Data
library(readxl)
Data <- read_excel("Data.xlsx")
View(Data)

library(mvtnorm)

Uji1=bartlett.test(Diameter~Batch,data=Data)
Uji1
```

Bartlett’s Test Using R

Contoh

Seorang petugas quality control ingin mengetahui homogenitas dari produk gear yang mereka hasilkan dari masing-masing batch, dimana dalam satu hari produksi terdapat 10 batch untuk proses produksi. Berdasarkan data yang telah disampel apakah sudah cukup membuktikan bahwa produk yang dihasilkan adalah homogen antar batch?

Lihat Dataset yang sudah diupload di HEBAT

Levene’s Test

Levene Test

- Levene’s test merupakan pengujian yang dikenalkan Levene pada tahun 1960.
- Uji ini digunakan untuk melihat kesamaan varians dari *k* sampel.
- Jika memiliki varians yang sama, maka dapat diartikan bahwa sampel tersebut bersifat homogen.
- Beberapa pengujian statistik mensyaratkan asumsi varians antar group atau sampel adalah sama. Oleh sebab itu levene test dapat menjadi salah satu alternatif untuk menguji asumsi tersebut.
- Levene test merupakan salah satu alternatif pengujian selain Bartlett test, khususnya untuk data yang tidak normal.
- Tetapi jika keyakinan terhadap data adalah normal, atau mendekati normal maka Bartlett test memiliki perfoma yang lebih baik.

Levene Test

Hipotesis

H0 : $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

H1 : minimal ada satu $\sigma_i^2 \neq \sigma_j^2$

Statistik Uji

$$W = \frac{(N-k)\sum_{i=1}^{k} N_i \left(\bar{Z}_{i.} - \bar{Z}_{..}\right)^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Z_{ij} - \bar{Z}_{i.}\right)^2}$$

Dimana,

N_i : jumlah sampel dalam group ke-*i*

N : jumlah sampel

k : jumlah group

$$Z_{ij} = \left|Y_{ij} - \bar{Y}_{i.}\right|$$

Mean

$$Z_{ij} = \left|Y_{ij} - \tilde{Y}_{i.}\right|$$

Median

Kriteria Penolakan

- Jika nilai statistik uji (W) lebih besar dari nilai tabel ($F_{\alpha,k-1,N-k}$), maka tolak H0.
- Jika nilai statistik uji lebih kecil dari nilai tabel, maka gagal tolak H0

Levene Test Using R

```
#Setting Directory
setwd("D:/UNAIR/1. Perkuliahan/Statnonpar/2023-2024/M4")

#Import Data
library(readxl)
Data <- read_excel("Data.xlsx")
View(Data)

library(misty)
Uji2=test.levene(Diameter~Batch,data=Data,method = "mean", conf.level = 0.95)
Uji2
```


Terima kasih

Statistika Non Parametrik
TSD - Ganjil 2023/2024

Pertemuan Minggu 5:
Uji Hipotesis 2 Sampel Independen & Dependen

Outline

Uji komparatif dua sampel independen

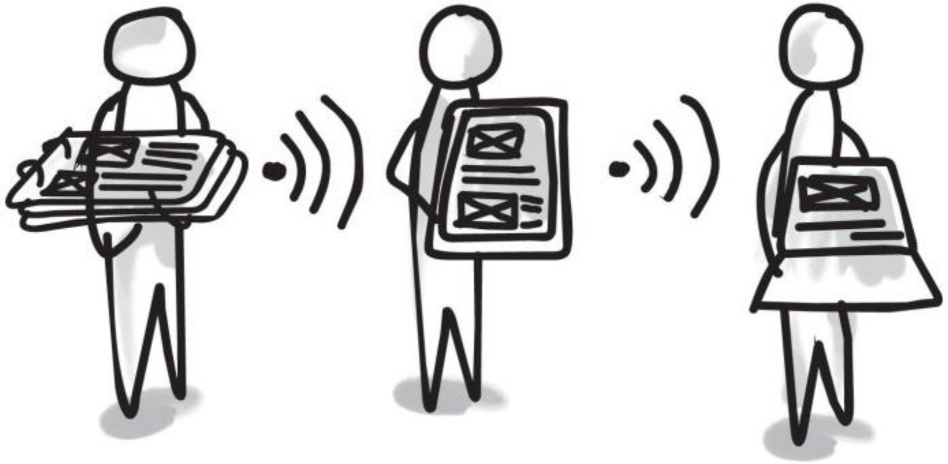
- uji median
- uji mann whitney

Uji komparatif dua sampel dependen

- mc nemar

Pendahuluan

- ✓ Uji dua sampel biasanya juga disebut uji komparasi dua sampel.
- ✓ Merupakan uji yang dilakukan untuk membandingkan parameter lokasi dari dua sampel yang berasal dari dua populasi.
- ✓ Uji ini terdiri dari dua bagian yaitu uji dua sampel Independen (bebas) dan Dependen (berhubungan).



Definisi dan Tujuan Uji Hipotesis 2 Sampel Independen

- ✓ Sampel bebas adalah dua kelompok sampel yang berbeda, namun dilakukan pengukuran pada waktu yang sama.
- ✓ Pengujian dua sampel bebas bertujuan untuk menganalisis perbedaan dari dua sampel yang saling bebas, atau dengan kata lain untuk menguji apakah kedua sampel yang saling bebas tersebut berasal dari populasi yang memiliki karakteristik yang sama atau tidak.

Pedoman Memilih Metode

Skala yang Digunakan	Alat Analisis Pengujian Hipotesis Dua Sampel Bebas
Nominal	Uji Fisher Exact Probability Uji Chi Square Two Sample
Ordinal	**Uji Median** **Uji Mann-Whitney (U Test)** Uji Kolmogorov-Smirnov Two Samples Uji Wald-Woldfowitz

Uji Median

Uji Median

- Merupakan uji untuk dua sampel independen yang paling sederhana.
- Digunakan untuk menguji apakah sama antara median dari sampel 1 dengan median dari sampel 2.

Asumsi yang Harus Dipenuhi :

- Kedua sampel diambil secara acak dari dua populasi independen
- Sampel I dengan pengamatan $X_1, X_2, \ldots, X_n$
- Sampel II dengan pengamatan $Y_1, Y_2, \ldots, Y_n$
- Variabel pengamatan keduanya adalah kontinyu
- Skala pengukuran minimal ordinal
- Bila median kedua populasi adalah sama maka peluang p (banyaknya pengamatan diatas median gabungan (*grand median*)) adalah sama untuk keduanya.

Uji Median

Prosedur:

- Tentukan **median gabungan** yaitu median untuk semua skor dalam kedua sampel.
- Pisahkan skor masing-masing kelompok berdasarkan median gabungan dan **masukan frekuensi-frekuensi** yang diperoleh **dalam tabel** (slide selanjutnya).
- Jika ada skor yang mempunyai **nilai sama dengan median gabungan**, peneliti mempunyai dua pilihan solusi: Pertama, jika kasus yang nilainya sama dengan median gabungan sedikit dan n gabungan besar maka nilai tersebut bisa tidak digunakan dalam analisis. Kedua, kasus yang nilainya sama dengan median gabungan dapat dibagi dua menjadi skor-skor di atas median dan skor-skor di bawah median.
- Jika $n \leq 20$ gunakan uji Fisher exact.
- Jika $n > 20$ gunakan uji Chi-square dengan koreksi kontinyuitas (jika frekuensi harapannya ada yang kurang dari 5, maka gunakan uji Fisher).

Uji Median

Struktur Data

	Kelompok I	Kelompok II	Jumlah
> median	A	B	A + B
< median	C	D	C + D
	n1 = A + C	n2 = B + D	n = n1 + n2

Keterangan :

A = banyaknya pengamatan > nilai median gabungan (*grand median*) dari kelompok I

B = banyaknya pengamatan > nilai median gabungan (*grand median*) dari kelompok II

C = banyaknya pengamatan < nilai median gabungan (*grand median*) dari kelompok I

D = banyaknya pengamatan < nilai median gabungan (*grand median*) dari kelompok II

n = banyak pengamatan keseluruhan

Uji Median

Hipotesis :

Pada uji median, hipotesis yang digunakan hanya dua sisi (karena proses satu sisi sangatlah rumit)

- H_0 : Kedua populasi memiliki median yang sama
- H_1 : Kedua populasi memiliki median yang berbeda

Statistik Uji :

Uji Fisher (n ≤ 20):

$$p = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{n!A!B!C!D!}$$

Tolak H_0 jika probabilitas (p) ≤ α.

Uji Median

Statistik Uji :

Uji Chi-square (n > 20):

$$\chi^2 = \frac{n\left(\left|AD - BC\right| - \frac{n}{2}\right)^2}{(A+B)(C+D)(A+C)(B+D)}, \text{ dengan df } = 1$$

Tolak H_0 jika $\chi^2_{hitung} \geq \chi^2_{tabel(\alpha,df)}$.

Statistik Uji :

Untuk sampel besar juga bisa menggunakan pendekatan normal dengan statistik uji sebagai berikut:

$$T = \frac{(A/n_1) - (B/n_2)}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Dengan : $\hat{P} = (A+B)/N$

Tolak H_0 jika $T \geq Z_{\alpha/2}$ atau $T \leq -Z_{\alpha/2}$

Uji Median

Contoh :

Untuk melihat apakah ada perbedaan produksi per-hektar tanaman padi dengan 2 metode penanaman yang berbeda, pertumbuhan tanaman padi dipilih dari sejumlah bidang tanah yang berbeda secara random. Kemudian produksi per-hektar dari masing-masing bidang tanah dihitung dan hasilnya adalah sebagai berikut: (α = 5%)

Metode 1	83	91	94	89	96	91	92	90	92	85		
Metode 2	91	90	81	83	84	83	88	91	90	84	80	85

Penyelesaian

Hipotesis:

H_0 : dua metode tanam mempunyai nilai median yang sama untuk produksi per hektar.

H_1 : dua metode tanam mempunyai nilai median yang berbeda untuk produksi per hektar.

Uji Median

Nilai median gabungan = 89.5

Karena n > 20 gunakan uji **Chi-square**

$$\chi^2 = \frac{n\left(|AD - BC| - \frac{n}{2}\right)^2}{(A+B)(C+D)(A+C)(B+D)}$$

$$\chi^2 = \frac{22\left(|56 - 12| - \frac{22}{2}\right)^2}{(11)(11)(10)(12)} = 1.65$$

$$\chi^2_{tabel} = \chi^2_{0.05,1} = 3.84$$

	Metode 1	Metode 2	Jumlah
> median	7	4	11
< median	3	8	11
	n1 = 10	n2 = 12	n = **22**

Keputusan:
Gagal tolak H_0 karena $\chi^2_{hitung} < \chi^2_{tabel(\alpha,df)}$.
Kesimpulan:
Kesimpulan: dua metode tanam mempunyai nilai median yang sama untuk produksi per-hektar dengan tingkat keyakinan 95%

Uji Median

Nilai median gabungan = 89.5

Menggunakan pendekatan **Normal**

$$\hat{P} = (A+B)/\,\mathrm{n} = 11/22 = 0.5$$

$$T = \frac{(7/10)-(4/12)}{\sqrt{0.5(1-0.5)\left(\frac{1}{10}+\frac{1}{12}\right)}} = 1.7127$$

$Z_{0.025} = 1.96$

Tolak H_0 jika $T \geq Z_{\alpha/2}$ atau $T \leq -Z_{\alpha/2}$

	Metode 1	Metode 2	Jumlah
> median	7	4	11
< median	3	8	11
	n1 = 10	n2 = 12	n = **22**

Keputusan:
Gagal tolak H_0 karena $T < Z_{\alpha/2}$.
Kesimpulan:
Kesimpulan: dua metode tanam mempunyai nilai median yang sama untuk produksi per-hektar dengan tingkat keyakinan 95%

Uji Mann-Whitney

Uji Mann-Whitney

- Uji mann Whitney ≈ Uji U mann Whitney
- Digunakan untuk menguji parameter lokasi dari 2 populasi bila skala datanya ordinal.

Asumsi:

- data merupakan sampel acak masing-masing dari dua populasi
- populasi 1 dengan pengamatan $X_1, X_2, \ldots\ldots\ldots, X_{n1}$
- populasi 2 dengan pengamatan $Y_1, Y_2, \ldots\ldots\ldots, Y_{n1}$
- kedua sampel independen
- variabel pengamatan adalah variabel acak kontinyu
- skala pengukuran minimal ordinal
- fungsi distribusi kedua populasi hanya berbeda dalam hal populasi lokasi

Uji Mann-Whitney

Perumusan Hipotesis :

- Uji Dua Sisi

  H_0 : Kedua populasi yang diamati identik

  H_1 : Kedua populasi yang diamati berbeda

- Satu Sisi (Arah Kiri)

  H_0 : Kedua populasi yang diamati memiliki distribusi yang identik

  H_1 : Nilai-nilai X cenderung lebih kecil daripada nilai Y

- Satu Sisi (Arah Kanan)

  H_0 : Kedua populasi yang diamati memiliki ditribusi yang identik

  H_1 : Nilai-nilai X cenderung lebih besar daripada nilai Y

Uji Mann-Whitney

Prosedur :

a. Gabungkan dua sampel independen dan beri jenjang pada tiap-tiap anggotanya mulai dari pengamatan terkecil sampai nilai pengamatan terbesar. Jika ada dua atau lebih pengamatan yang sama maka digunakan jenjang rata-rata.

b. Hitunglah jumlah jenjang masing-masing bagi sampel pertama dan kedua dan beri notasi R1 dan R2.

c. Untuk pengujian statistik U, kemudian dihitung dari sampel pertama dengan n1 pengamatan.

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

Atau dari sampel kedua dengan n_2 pengamatan

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

Uji Mann-Whitney

d. Dari dua nilai U tersebut yang digunakan adalah nilai U yang lebih kecil. Untuk dibandingkan dengan nilai U tabel.

e. Bandingkan nilai U dengan nilai U dalam tabel (untuk n_1 dan n_2 yang lebih kecil dari 20.

f. Pengambilan keputusan dengan kriteria:

- H_0 tidak dapat ditolak, jika $U \geq U_{\alpha, n1, n2.}$
- H_0 ditolak, jika $U < U_{\alpha, n1, n2.}$

Uji Mann-Whitney

Aproksimasi Sampel Besar :

- Bila n_1 atau n_2 > 20 maka diterapkan:

$$Z = \frac{U - \mu_U}{\sigma_U}, \quad \mu_U = \frac{n_1 n_2}{2}$$

U = Nilai terkecil antara U_1 dan U_2

- Bila tidak ada angka sama

$$\sigma_U = \frac{\sqrt{n_1 n_2 (n_1 + n_2 + 1)}}{12}$$

- Bila ada angka-angka sama dalam kelompok yang berbeda, dilakukan koreksi:

$$\sigma_U = \sqrt{\left(\frac{n_1 n_2}{n(n-1)}\right)\left(\frac{n^3 - n}{12} - \sum T\right)}$$

T = banyak observasi berangka sama

$$T = \frac{t^3 + t}{12}; \; n = n_1 + n_2$$

Daerah Penolakan

- Dua sisi, tolak H_0 jika

  $Z \geq Z_{\alpha/2}$ atau $Z \leq Z_{1-\alpha/2}$
- Satu sisi (arah kiri)

  $Z \leq Z_{1-\alpha}$
- Satu sisi (arah kanan)

  $Z \geq Z_{\alpha}$

Contoh

Seorang manajer pemasaran sirup, melakukan penelitian dengan tujuan untuk menguji apakah perbedaaan selera konsumen sirup Rasa Durian di dua desa yaitu di Desa Karanganyar dan di Desa Kali Tengah. Untuk kepentingan tersebut diambil sampel secara acak sebanyak 14 konsumen dari dua desa tersebut dengan data sebagai berikut:

Desa	Sikap
Karanganyar	Sangat Suka
Karanganyar	Suka
Karanganyar	Sangat Suka
Karanganyar	Suka
Karanganyar	Sangat Suka
Karanganyar	Cukup Suka
Karanganyar	Tidak Suka
Kalitengah	Suka
Kalitengah	Suka
Kalitengah	Sangat Suka
Kalitengah	Suka
Kalitengah	Tidak Suka
Kalitengah	Suka
Kalitengah	Cukup Suka
Kalitengah	Sangat Suka
Kalitengah	Sangat Suka

Penyelesaian

1. Judul Penelitian

Perbedaan Selera Konsumen terhadap Sirup Rasa Durian di Desa Karanganyar dan di Desa Kalitengah.

2. Variabel Penelitian

Selera konsumen terhadap sirup rasa durian di Desa Karanganyar dan Selera konsumen sirup rasa durian di Kalitengah

3. Pertanyaan Penelitian

Apakah terdapat perbedaan selera konsumen terhadap sirup rasa durian di Desa Karanganyar dan di Desa Kalitengah?

4. Hipotesis

H_0 : Tidak terdapat perbedaan selera konsumen terhadap sirup rasa durian di Desa Karanganyar dan di Desa Kalitengah.

H_1 : Terdapat perbedaan selera konsumen terhadap sirup rasa durian di Desa Karanganyar dan di Desa Kalitengah.

atau:

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Penyelesaian

5. Kriteria Pengujian

H_0 tidak dapat ditolak, jika $U \geq U\alpha$, atau Sig. > 0.05

Ho ditolak, jika $U < U\alpha$, atau Sig. ≤ 0.05

No	Sikap Masyarakat Karanganyar	Ranking A	Sikap Masyarakat Kalitengah	Ranking B
1	5	13.5	4	7.5
2	4	7.5	4	7.5
3	5	13.5	5	13.5
4	4	7.5	4	7.5
5	5	13.5	2	1.5
6	3	3.5	4	7.5
7	2	1.5	3	3.5
8			5	13.5
9			5	13.5
		60.5		75.5

Penyelesaian

$R_1 = 60.5$ dan $R_2 = 75.5$.

$$\boxed{U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1}$$

$$\boxed{U_1 = 7x9 + \frac{7(7+1)}{2} - 60{,}5 = 30{,}5}$$

$$\boxed{U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2}$$

$$\boxed{U_2 = 7x9 + \frac{9(9+1)}{2} - 75{,}5 = 32{,}5}$$

Output SPSS

Ranks

	kota	N	Mean Rank	Sum of Ranks
sikap	Karanganyar	7	8.64	60.50
	Kalitengah	9	8.39	75.50
	Total	16		

Test Statistics[b]

	sikap
Mann-Whitney U	30.500
Wilcoxon W	75.500
Z	-.112
Asymp. Sig. (2-tailed)	.911
Exact Sig. [2*(1-tailed Sig.)]	.918[a]

a. Not corrected for ties.

b. Grouping Variable: kota

Kesimpulan

- Berdasarkan analisis di atas ternyata $U_1 < U_2$, sehingga yang digunakan untuk membandingkan dengan U tabel adalah U_1.
- Nilai U_{tabel} pada Tabel Mann- Whitney U test, dengan $\alpha = 0.05$; n1 = 7; n2 = 9, adalah $U_{tabel} = 9$.
- Karena U_{hitung} (30.5) > U_{tabel} (9), atau. Sig. 2-tail (0.911) > alpha (0.05), maka hipotesis nol tidak dapat ditolak.
- Sehingga hipotesis yang menyatakan ”Terdapat perbedaan selera konsumen sirup rasa durian di desa Karanganyar dan Kalitengah”, ditolak.

Mc Nemar Test

Mc Nemar Test

- ✓ Merupakan salah satu uji untuk 2 sampel berhubungan khususnya jika data berupa frekuensi.
- ✓ Pada uji ini ada 2 kategori pengamatan yang salah satunya diperhatikan, selain dari itu tidak.
- ✓ Data dapat dikategorikan sebagai peristiwa ”ya” atau ”tidak”

Asumsi :

- Data terdiri dari N subyek atau berupa data berpasangan atau berasal dari dua sampel yang dipasangkan.
- Skala pengukuran adalah nominal, dengan peristiwa hasil amatan terhadap 2 sampel yang berhubungan berupa ”ya-ya”, ”ya-tidak”, ”tidak-ya”, dan ”tidak-tidak”.
- Jika berupa 2 sampel maka salah satu sampel merupakan kontrol dari sampel lainnya yang dalam hal ini sebagai eksperimennya.

Mc Nemar Test

- Struktur data hasil pengamatan uji Mc Nemar berupa tabel kontingensi 2 x 2

		Sampel 1		Jumlah
		Ya	Tidak	
Sampel 2	Ya	A	B	A + B
	Tidak	C	D	C + D
Jumlah		A + C	B + D	N

- Dalam uji Mc Nemar ini ingin diketahui apakah ada perbedaan dari suatu kategori, sehingga pengujian hipotesis yang diujikan berupa pengujian 2 arah.

Mc Nemar Test

Hipotesis :

- $H_0 : P_1 = P_2$
- $H_1 : P_1 \neq P_2$

Atau

- $H_0 : P_1 - P_2 = 0$
- $H_1 : P_1 - P_2 \neq 0$

Dengan P_1 adalah proporsi kategori amatan yang diperhatikan dan P_2 adalah proporsi kategori lainnya.

- Untuk pengamatan sampel maka :

$$\hat{p}_1 = \frac{A+B}{N} \text{ dan } \hat{p}_2 = \frac{A+C}{N}$$

- Karakteristik yang diperhatikan adalah "Ya" maka selisih :

$$\hat{p}_1 - \hat{p}_2 = \frac{B-C}{N}$$

- Statistik uji :

$$Z = \frac{B-C}{\sqrt{B+C}}$$

- Daerah kritis :

Tolak H_0 jika $|Z| > Z_{1-\alpha/2}$

Mc Nemar Test

Contoh:

Pike dan Smith menggunakan data yang dikumpulkan oleh Johnson dan Johnson untuk menjelaskan uji McNemar. Mereka memasangkan masing-masing dari 85 pasien yang dirawat karena menderita penyakit Hodgkin dengan saudara kandung mereka sendiri yang sehat, berjenis kelamin sama, dan berbeda usia tidak lebih dari 5 tahun (sebagai kontrol). Apakah ada perbedaan dalam sejarah tonsilektomi pada kedua kelompok?

		Tonsilektomi kontrol		Jumlah
		Ya	Tidak	
Tonsilektomi pasien	Ya	26	15	41
	Tidak	7	37	44
Jumlah		33	52	85

Penyelesaian

Hipotesis

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

Statistik uji

$$z = \frac{15 - 7}{\sqrt{15 + 7}} = 1{,}71$$

Keputusan:

Nilai 1,71 lebih kecil daripada 1,96 sehingga kita tidak dapat menolak H_0. Oleh karena ini adalah uji dua sisi, nilai *p* untuk contoh ini adalah 2(0,0436) = 0,0872 > 0,05. Jadi, tidak ada perbedaan dalam sejarah tonsilektomi pada kedua kelompok.

Latihan soal

Latihan 1

Seorang pengusaha ingin mengetahui frekuensi belanja online dan belanja offline tiap bulan (datang langsung ke supermarket / pasar) masyarakat yang tinggal di kota Surabaya. Diambil sampel beberapa warga dengan frekuensi belanja online maupun offine adalah sebagai berikut. Apakah terdapat perbedaan frekuensi belanja online dan offline untuk warga Surabaya? Gunakan uji median dengan alpha = 1%.

Frekuensi Belanja online	Frekuensi Belanja offline
5	3
1	2
10	10
3	3
6	3
16	2
3	
4	
9	
10	

Latihan 2

Dilakukan penelitian untuk mengetahui adakah perbedaan keuntungan antara penambang bitcoin dengan ethereum. Penelitian menggunakan sampel 17 orang yang mempunyai aset bitcoin dan 23 orang yang mempunyai asset ethereum. Selanjutnya kedua kelompok jenis asset tersebut diukur besaran keuntungannya. Gunakan uji Mann-Whitney untuk mengetahui perbedaan kualitas antara kedua jenis kripto tersebut dengan alpha = 1%.

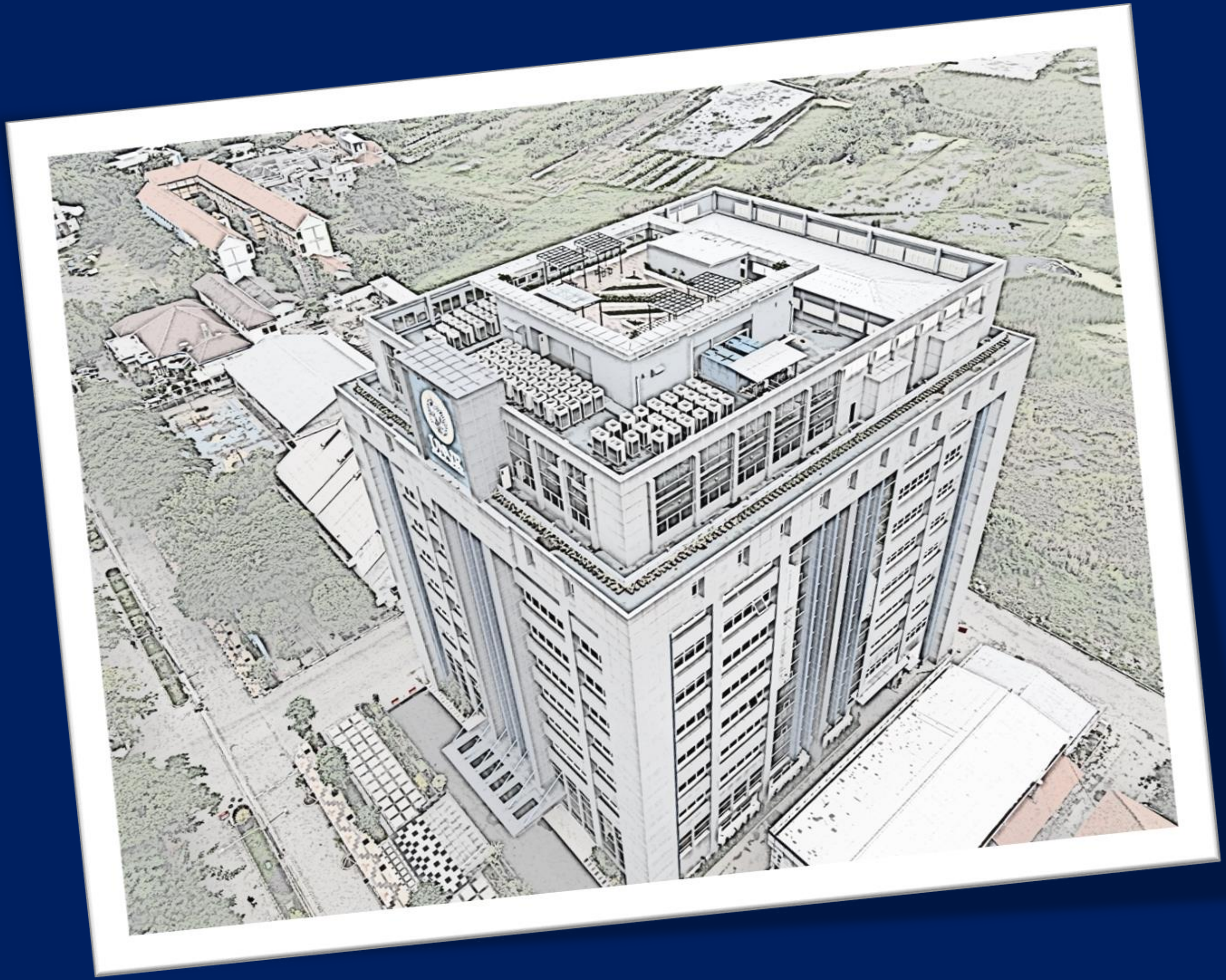
Keuntungan bitcoin	Keuntungan Ethereum
101	191
183	262
219	274
196	282
264	210
240	235
271	212
242	242
248	258
174	295
266	260
169	261
200	290
301	190
450	89
141	101
98	250
	78
	98
	70
	103
	77
	303

Latihan 3

Waters melaksanakan suatu uji klinik terkontrol untuk menangani sakit kepala sebelah (migraine headache) di kalangan wanita. Untuk itu, sejumlah subyek menerima tablettablet ergotamine (1 mg) dan suatu placebo (laktosa) dengan urutan pemberian yang acak selama periode delapan minggu. Hasil eksperimen seperti dalam tabel berikut.

		Bermanfaatkah ergotamine?		Jumlah
		Ya	**Tidak**	
Bermanfaatkah placebo?	**Ya**	22	24	46
	Tidak	18	15	33
Jumlah		40	39	79

Berdasarkan data di atas, dengan menggunakan Mc Nemar Test, dapatkah kita menyimpulkan bahwa terapi eksperimental tersebut efektif untuk menangani sakit kepala sebelah? (Misalkan taraf nyata sebesar 5%)



Terima Kasih

Critical Values of the Mann-Whitney U
(Two-Tailed Testing)

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	.05	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	.01	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	.05	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	.01	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	.05	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
	.01	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	.05	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	.01	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	.05	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	.01	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	.05	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
	.01	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	.05	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	.01	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	.05	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	.01	3	7	12	17	22	28	33	39	45	51	56	63	69	74	81	87	93	99
20	.05	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	.01	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Critical Values of the Mann-Whitney U

(One-Tailed Testing)

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
	.01	--	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	.01	--	--	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	.05	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	.01	--	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	.05	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	.01	--	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	.05	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
	.01	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	.05	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
	.01	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	.05	4	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
	.01	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	.05	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
	.01	1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11	.05	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
	.01	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	.05	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
	.01	2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	.05	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
	.01	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	.05	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
	.01	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	.05	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
	.01	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	.05	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
	.01	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	.05	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
	.01	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	.05	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
	.01	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	.05	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	.01	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	.05	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	.01	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Statistika Non Parametrik
TSD - Ganjil 2022/2023

Pertemuan 5 :

Uji Keselarasan / Kesesuaian / GoF

Outline

1. Uji Chi-square
2. Uji Kolmogorov-Smirnov
3. Uji distribusi Uniform (dengan uji Chi-Sq)
4. Uji distribusi Binomial (dengan uji Chi-Sq)
5. Uji distribusi Poisson (dengan uji Chi-Sq)
6. Uji distribusi Uniform (dengan uji KS) (Next Week)
7. Uji distribusi Binomial (dengan uji KS) (Next Week)
8. Uji distribusi Poisson (dengan uji KS) (Next Week)
9. Uji distribusi Normal (dengan uji Chi-Sq, KS, Liliefors) (Next Week)

UJI KESELARASAN/ KESESUAIAN/ GOODNES OF FIT

- Ada dua macam cara pengujian untuk keselarasan/kesesuaian data dengan suatu distribusi peluang tertentu yaitu dengan menggunakan **UJI CHI-SQUARE** dan **UJI KOLMOGOROV SMIRNOV.** Sedangkan khusus untuk menguji keselarasan distribusi normal masih ada uji lainnya yaitu **UJI LILLIEFORS**.
- Pengujian hipotesis keselarasan (*goodness of fit*) merupakan pengujian hipotesis untuk menentukan apakah suatu himpunan frekuensi yang diharapkan sama dengan frekuensi yang diperoleh dari suatu distribusi, seperti distribusi binomial, poisson, normal, atau dari perbandingan lain.
- Jadi, uji *goodness of fit* merupakan pengujian kecocokan atau kebaikan antara hasil pengamatan (frekuensi pengamatan) tertentu dengan frekuensi yang diperoleh berdasarkan nilai harapannya (frekuensi teoretis)

UJI KESELARASAN/ KESESUAIAN/ GOODNES OF FIT

- ASUMSI – ASUMSI :
  1. sampel terdiri n pengamatan bebas
  2. skala pengukuran minimal yang mungkin digunakan nominal
  3. hasil pengamatan diklasifikasikan dalam r kategori yang tidak saling tumpang tindih
- Adapun bentuk perumusan hipotesisnya secara umum dapat dituliskan seperti berikut di bawah ini.
- HIPOTESIS:

  H_0 : data sampel berasal dari suatu populasi berdistribusi ....

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi ....

Uji Chi-square

Uji Chi-square

Langkah-langkah yang dilakukan secara umum dalam pengujian chi square sebagai berikut :

a. Membuat formulasi hipotesis (H_0 dan H_1)

b. Menentukan taraf signifikansi (α)

c. Memilih uji statistik yang sesuai (χ^2_{hitung} dibandingkan dengan χ^2_{tabel})

d. Menentukan kesimpulan / pengambilan keputusan

Uji Chi-square

Perumusan hipotesis

H_0 : **tidak ada** perbedaan antara frekuensi yang diamati dengan frekuensi yang diharapkan
H_1 : **ada** perbedaan antara frekuensi yang diamati dengan frekuensi yang diharapkan

Statistik uji

$$\chi^2_{hitung} = \sum \left[\frac{(f_o - f_e)^2}{f_e} \right]$$

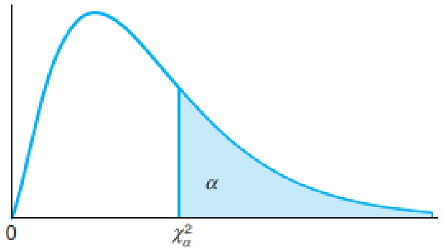
f_o = besarnya frekuensi yang diamati (*observed*)

f_e = besarnya frekuensi yang diharapkan (*expected*)

Daerah penolakan

Tolak H_0 apabila $\chi^2_{hitung} \geq \chi^2_{tabel} \longrightarrow \chi^2_{\alpha,(k-1)}$

k : banyaknya kategori atau jumlah kelas



Contoh 1

Untuk menarik konsumen dilakukan pembungkusan barang dengan menggunakan warna yang berbeda. Dari pasaran bebas diteliti pilihan warna dari konsumen. Hasilnya dari 1000 barang ternyata para konsumen telah membeli dengan pembungkus warna merah, hijau, biru, dan kuning berturut-turut 205, 286, 315 dan 194. Apakah penelitian ini berhasil memperlihatkan bahwa warna-warna pembungkus berlainan telah mengakibatkan selera pembeli yang berlainan pula? Gunakan taraf signifikansi 5% .

Jawab

- **Perumusan hipotesis**

H_0 : warna pembungkus **tidak mempengaruhi** selera pembeli

H_1 : warna pembungkus **mempengaruhi** selera pembeli

- **Taraf Signifikansi**

$\alpha = 5\%$

- **Statistik uji**

Warna	fo	fe	(fo-fe)	(fo-fe)^2	(fo-fe)^2/fe
Merah	205	250	-45	2025	8.1
Hijau	286	250	36	1296	5.184
Biru	315	250	65	4225	16.9
Kuning	194	250	-56	3136	12.544
Jumlah	**1000**	**1000**			**42.728**

$k = 4$

$$\chi^2_{hitung} = \sum\left[\frac{(f_o - f_e)^2}{f_e}\right] = 42{,}728$$

- **Daerah penolakan**

Berdasarkan Tabel diperoleh $\chi^2_{tabel} = \chi^2_{\alpha,(k-1)} = \chi^2_{5\%,(4-1)} = \chi^2_{5\%,(3)} = \boxed{7{,}815}$

Daerah kritisnya adalah $\geq 7{,}815$

- **Kesimpulan**

Karena $\chi^2_{hitung} > \chi^2_{tabel}$ maka TOLAK H_0

Artinya penelitian ini berhasil memperlihatkan bahwa **warna-warna pembungkus berlainan** telah mengakibatkan **selera pembeli yang berlainan** pula

Table A.5 (continued) Critical Values of the Chi-Squared Distribution

	α						
v	**0.30**	**0.25**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668
5	6.064	6.626	7.289	9.236	11.070	12.832	13.388

Contoh 2

Bagaimana hubungan antara **pengetahuan**, **sikap** dan **perilaku** wanita hamil tentang mual-muntah di poliklinik kebidanan RSCM tahun 2021?

Kerangka konsep penelitian: Pengetahuan→Sikap→Perilaku

Dikelompokkan menjadi **Baik, Cukup,** dan **Kurang**

Ujilah apakah ada hubungan antara **Pengetahuan** dan **Perilaku** wanita hamil di poliklinik kebidanan RSCM tahun 2021.

	PERILAKU			
PENGETAHUAN	**Baik**	**Cukup**	**Kurang**	
Baik	18 (a)	10 (b)	7 (c)	**35**
Cukup	13 (d)	14 (e)	13 (f)	**40**
Kurang	13 (g)	12 (h)	25 (i)	**50**
TOTAL	**44**	**36**	**45**	**125**

Jawab

- **Perumusan hipotesis**

H_0 : Pengetahuan **tidak mempengaruhi** perilaku ibu hamil

H_1 : Pengetahuan **mempengaruhi** perilaku ibu hamil

- **Statistik uji**

Pengetahuan **baik**, perilaku **baik** (sel a)

- observed=18
- Expected=$\frac{total\ baris\ 1 \times total\ kolom\ 1}{grand\ total} = \frac{35 \times 44}{125} = 12{,}32$

Sel	Observed	Expected	(O-E)	(O-E)^2	(O-E)^2/E
a	18	12.32	5.68	32.2624	2.618701
b	10	10.08	-0.08	0.0064	0.000635
c	7	12.6	-5.6	31.36	2.488889
d	13	14.08	-1.08	1.1664	0.082841
e	14	11.52	2.48	6.1504	0.533889
f	13	14.4	-1.4	1.96	0.136111
g	13	17.6	-4.6	21.16	1.202273
h	12	14.4	-2.4	5.76	0.4
i	25	18	7	49	2.722222
Total	**125**	**125**			**10.18556**

$$\chi^2_{hitung} = \sum\left[\frac{(f_o - f_e)^2}{f_e}\right] = 10{,}185$$

- **Taraf Signifikansi**

$\alpha = 5\%$

- **Daerah penolakan**

df=(baris-1)(kolom-1)

Berdasarkan Tabel diperoleh $\chi^2_{tabel} = \chi^2_{\alpha,(3-1)(3-1)} = \chi^2_{5\%,(4)} = \boxed{9{,}488}$

Daerah kritisnya adalah $\geq 9{,}488$

- **Kesimpulan**

Karena $\chi^2_{hitung} > \chi^2_{tabel}$ maka TOLAK H_0

Artinya penelitian ini berhasil memperlihatkan bahwa **pengetahuan mempengaruhi perilaku ibu hamil**

Table A.5 (continued) Critical Values of the Chi-Squared Distribution

	α						
v	**0.30**	**0.25**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668
5	6.064	6.626	7.289	9.236	11.070	12.832	13.388

Uji Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov

K-S 1 variable digunakan untuk membandingkan distribusi pengamatan dengan distribusi teoritis pada 1 variabel dengan skala ordinal

K-S 2 variabel digunakan untuk mencari sebab dan akibat berbeda dari 2 variabel dengan skala ordinal

$$KS_{\text{hitung}} = \mathbf{max}\left|D\right|$$

$$D = F_{\text{observed}} - F_{\text{expected}}$$

Daerah penolakan

Tolak H_0 apabila $KS_{\text{hitung}} > KS_{\text{tabel}}$

Contoh 3

Rasa sakit pada saat melahirkan ditunjukkan dengan nilai skor (1-5) oleh 10 orang wanita. Tunjukkan apakah ada perbedaan dalam pemilihan skor rasa sakit. Dengan taraf signifikansi 5%.

Skor	1	2	3	4	5	**Jumlah**
Jumlah Ibu	0	1	0	5	4	**10**

Jawab

- **Perumusan hipotesis**

H_0 : Tidak ada perbedaan dalam pemilihan skor rasa sakit

H_1 : Terdapat perbedaan dalam pemilihan skor rasa sakit

- **Statistik uji**

Skor	1	2	3	4	5	Jumlah
Jumlah Ibu	0	1	0	5	4	**10**
p observed	0	0.1	0	0.5	0.4	
p expected	0.2	0.2	0.2	0.2	0.2	
F Observed	0	0.1	0.1	0.6	1	
F Expected	0.2	0.4	0.6	0.8	1	
(FO-FE)	-0.2	-0.3	-0.5	-0.2	0	
|FO-FE|	0.2	0.3	**0.5**	0.2	0	
			KS hitung			

$$KS_{\text{hitung}} = \mathbf{max}|D| = 0,5$$

- **Taraf Signifikansi**

$\alpha = 5\%$

- **Daerah penolakan**

Berdasarkan Tabel diperoleh

$$KS_{\text{tabel}} = \boxed{0,409}$$

- **Kesimpulan**

Karena $KS_{\text{hitung}} > KS_{\text{tabel}}$ maka TOLAK H_0

Artinya penelitian ini berhasil memperlihatkan bahwa **terdapat perbedaan dalam pemilihan skor rasa sakit**

Tabel Nilai Kritis Uji Kolmogorov-Smirnov

n	α = 0,20	α = 0,10	α = 0,05	α = 0,02	α = 0,01
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432

Uji distribusi Uniform (dengan Chi-sq Test)

Uji distribusi Uniform (dengan Chi-sq Test)

- **Perumusan hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Uniform

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Uniform

- **Statistik uji**

$$\chi^2 = \sum_{i=1}^{r} \frac{(o_i - e_i)^2}{e_i}$$

- **Daerah penolakan**

  Tolak H_0 apabila $\chi^2_{hitung} \geq \chi^2_{\alpha,(k-1)}$

- Pada pengujian keselarasan untuk distribusi Uniform, nilai peluangnya adalah $p_i = \frac{1}{r}$, i = 1, 2, ... , r. Sehingga $e_i = p_i \times n$, dengan r adalah banyaknya kelas dan n adalah banyaknya data observasi.

Contoh 4

Diketahui data pada table di bawah. Apakah sebaran nilai tersebut Uniform? Gunakan alpha = 0.05.

Nilai	A	B	C	D	E
Frekuensi	14	18	32	20	16

Jawab

- **Hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Uniform

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Uniform

- **Statistik uji**

  r = banyaknya karakteristik yang diamati

  $$p_i = \frac{1}{r} = \frac{1}{5} \quad ; \quad e_i = p_i \times n = \frac{1}{5} \times 100 = 20$$

  $$\boxed{\chi^2_{hitung} = \sum_{i=1}^{r} \frac{(o_i - e_i)^2}{e_i} = 10}$$

Nilai	A	B	C	D	E	Total
oi	14	18	32	20	16	**100**
pi	0.2	0.2	0.2	0.2	0.2	**1**
ei	20	20	20	20	20	**100**
(oi-ei)	-6	-2	12	0	-4	
(oi-ei)^2	36	4	144	0	16	
(oi-ei)^2/ei	1.8	0.2	7.2	0	0.8	**10**

- **Daerah penolakan**

  Tolak H_0 karena $\chi^2_{0.05,(5-1)} = 9,488$; $\chi^2 \geq \chi^2_{\alpha,(4)} \rightarrow 10 \geq 9,488$

- **Kesimpulan** : Data sampel tidak berdistribusi Uniform

Table A.5 (continued) Critical Values of the Chi-Squared Distribution

	α									
v	**0.30**	**0.25**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.001**
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412	6.635	7.879	10.827
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824	9.210	10.597	13.815
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.266
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668	13.277	14.860	18.466
5	6.064	6.626	7.289	9.236	11.070	12.832	13.388	15.086	16.750	20.515

Uji distribusi Binomial (dengan Chi-sq Test)

Uji distribusi Binomial

$$Pi = P(x = i) = \begin{pmatrix} r \\ I \end{pmatrix} p^{(i)} q^{(r-i)}$$

p = peluang sukses terjadi

$q = 1 - p$

- **Perumusan hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Binomial

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Binomial

Uji distribusi Binomial

- **Statistik uji**

$$\chi^2 = \frac{\sum_{i=1}^{r} (o_i - e_i)^2}{e_i}$$

- **Daerah penolakan**

  Tolak H_0 apabila $\chi^2 \geq \chi^2_{\alpha,(k-1)}$

Contoh 5

Uji kerapuhan 280 batang nylon ditekukkan pada 5 titik dan dicatat banyaknya patahan (0, 1, 2, 3, 4, 5). Uji apakah data berasal dari populasi berdistribusi Binomial dengan parameter $p = 0{,}5$.

Banyak patahan	**0**	**1**	**2**	**3**	**4**	**5**	**TOTAL**
Banyak batang nylon	157	69	35	17	1	1	280

Uji distribusi Poisson (dengan Chi-sq Test)

Uji distribusi Poisson

$$P_i = P(x = i) = \frac{e^{-\lambda}\lambda^x}{x!}, \lambda \approx \hat{\lambda} = rata - rata$$

- **Perumusan hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Poisson

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Poisson

Contoh 6

- Banyak pasien di ruang tunggu dalam kurun waktu atau interval per 30 detik. Uji apakah banyak pasien menunggu dalam interval per 30 detik mengikuti distribusi Poisson dengan parameter $\lambda = 3$?

Banyak pasien teramati	0	1	2	3	4	5	6	7	Total
Banyak interval	20	54	74	67	45	25	11	4	300

Terima Kasih

Statistika Non Parametrik
TSD - Ganjil 2022/2023

Pertemuan 6 :

Uji Keselarasan / Kesesuaian / GoF

Outline

1. Uji Chi-square
2. Uji Kolmogorov-Smirnov
3. Uji distribusi Uniform (dengan uji Chi-Sq)
4. Uji distribusi Binomial (dengan uji Chi-Sq)
5. Uji distribusi Poisson (dengan uji Chi-Sq)
6. Uji distribusi Uniform (dengan uji KS) (This Week)
7. Uji distribusi Binomial (dengan uji KS) (This Week)
8. Uji distribusi Poisson (dengan uji KS) (This Week)
9. Uji distribusi Normal (dengan uji Chi-Sq, KS, Liliefors) (This Week)

Uji Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov

K-S 1 variable digunakan untuk membandingkan distribusi pengamatan dengan distribusi teoritis pada 1 variabel dengan skala ordinal

K-S 2 variabel digunakan untuk mencari sebab dan akibat berbeda dari 2 variabel dengan skala ordinal

Jika X adalah variable yang akan diuji dengan menggunakan UJI KS, maka terlebih dahulu harus diurutkan dari yang paling kecil hingga paling besar.

$$KS_{\text{hitung}} = \mathbf{max}\left|D\right|$$

$$D = F_{\text{observed}} - F_{\text{expected}}$$

Daerah penolakan

Tolak H_0 apabila $KS_{\text{hitung}} > KS_{\text{tabel}}$

Contoh 3

Rasa sakit pada saat melahirkan ditunjukkan dengan nilai skor (1-5) oleh 10 orang wanita. Tunjukkan apakah ada perbedaan dalam pemilihan skor rasa sakit. Dengan taraf signifikansi 5%.

Skor	1	2	3	4	5	**Jumlah**
Jumlah Ibu	0	1	0	5	4	**10**

Jawab

- **Perumusan hipotesis**

H_0 : Tidak ada perbedaan dalam pemilihan skor rasa sakit

H_1 : Terdapat perbedaan dalam pemilihan skor rasa sakit

- **Statistik uji**

Skor	1	2	3	4	5	Jumlah
Jumlah Ibu	0	1	0	5	4	**10**
p observed	0	0.1	0	0.5	0.4	**1**
p expected	0.2	0.2	0.2	0.2	0.2	**1**
F Observed	0	0.1	0.1	0.6	1	
F Expected	0.2	0.4	0.6	0.8	1	
(Fo-Fe)	-0.2	-0.3	-0.5	-0.2	0	
|Fo-Fe|	0.2	0.3	**0.5**	0.2	0	
			KS hitung			

$$KS_{\text{hitung}} = \mathbf{max}|D| = 0,5$$

- **Taraf Signifikansi**

$\alpha = 5\%$

- **Daerah penolakan**

Berdasarkan Tabel diperoleh

$$KS_{\text{tabel}} = \boxed{0,409}$$

- **Kesimpulan**

Karena $KS_{\text{hitung}} > KS_{\text{tabel}}$ maka TOLAK H_0

Artinya penelitian ini berhasil memperlihatkan bahwa **terdapat perbedaan dalam pemilihan skor rasa sakit**

Tabel Nilai Kritis Uji Kolmogorov-Smirnov

n	α = 0,20	α = 0,10	α = 0,05	α = 0,02	α = 0,01
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432

Uji distribusi Uniform (dengan KS)

Uji distribusi Uniform (dengan KS)

- **Perumusan hipotesis**
  H_0 : data sampel berasal dari suatu populasi berdistribusi Uniform
  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Uniform

- **Statistik uji**

$$KS_{\text{hitung}} = \mathbf{max}\left|D\right|$$

- **Daerah penolakan**
  Tolak H_0 apabila $KS_{\text{hitung}} > KS_{\text{tabel}}$

- Pada pengujian keselarasan untuk distribusi Uniform, nilai peluangnya adalah $p_i = \frac{1}{r}$, i = 1, 2, ... , r.

Contoh 4 (KS)

Diketahui data pada table di bawah. Apakah sebaran nilai tersebut Uniform? Gunakan alpha = 0.05.

Nilai	A	B	C	D	E
Frekuensi	14	18	32	20	16

Jawab

- **Hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Uniform

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Uniform

- **Statistik uji**

  r = banyaknya karakteristik yang diamati

  $$p_i = \frac{1}{r} = \frac{1}{5}$$

  $$KS_{\text{hitung}} = \mathbf{max}|D| = 0,08$$

- **Daerah penolakan**

  Gagal tolak H_0 karena $KS_{\text{hitung}} = 0,08 < KS_{\text{tabel}} = 0,134$

- **Kesimpulan** : Data sampel berdistribusi Uniform

Uniform						
Nilai	**A**	**B**	**C**	**D**	**E**	**Total**
Frekuensi	14	18	32	20	16	**100**
p observed	0.14	0.18	0.32	0.2	0.16	**1**
p expected	0.2	0.2	0.2	0.2	0.2	**1**
F observed	0.14	0.32	0.64	0.84	1	
F expected	0.2	0.4	0.6	0.8	1	
F obs - F exp	-0.06	-0.08	0.04	0.04	0	
|F obs - F exp|	0.06	**0.08**	0.04	0.04	0	
		KS Hitung				

Tabel Nilai Kritis Uji Kolmogorov-Smirnov

n	α = 0,20	α = 0,10	α = 0,05	α = 0,02	α = 0,01
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
95	0,108	0,124	0,137	0,154	0,165
100	0,106	0,121	$\boxed{0,134}$	0,150	0,161

Uji distribusi Binomial (dengan KS)

Uji distribusi Binomial

$$Pi = P(x = i) = \begin{pmatrix} r \\ I \end{pmatrix} p^{(i)} q^{(r-i)}$$

p = peluang sukses terjadi

$q = 1 - p$

- **Perumusan hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Binomial

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Binomial

Uji distribusi Binomial

- **Statistik uji**

$$KS_{\text{hitung}} = \mathbf{max}\left|D\right|$$

- **Daerah penolakan**

Tolak H_0 apabila $KS_{\text{hitung}} > KS_{\text{tabel}}$

Contoh 7

Uji kerapuhan 280 batang nylon ditekukkan pada 5 titik dan dicatat banyaknya patahan (0, 1, 2, 3, 4, 5). Uji apakah data berasal dari populasi berdistribusi Binomial dengan parameter $p = 0{,}5$.

Banyak patahan	0	1	2	3	4	5	TOTAL
Banyak batang nylon	2	5	3	2	1	2	15

Jawab

- **Hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Binomial

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Binomial

- **Statistik uji**

  r = banyaknya karakteristik yang diamati

  $$KS_{\text{hitung}} = \mathbf{max}|D| = 0,268$$

Binomial							
Banyak patahan (i)	**0**	**1**	**2**	**3**	**4**	**5**	**TOTAL**
Banyak Batang Nylon	**2**	**5**	**3**	**2**	**1**	**2**	**15**
p observed	0.133333	0.333333	0.2	0.133333	0.066667	0.133333	**1**
F observed	0.133333	0.466667	0.666667	0.8	0.866667	1	
combin	5	1	10	1	10	5	
p^i	1	0.5	0.25	0.125	0.0625	0.03125	
q^(r-i)	0.03125	0.0625	0.125	0.25	0.5	1	
p expected	0.15625	0.03125	0.3125	0.03125	0.3125	0.15625	**1**
F expected	0.15625	0.1875	0.5	0.53125	0.84375	1	
|Fo-Fe|	0.022917	0.279167	0.166667	**0.26875**	0.022917	0	
				KS hitung			

- **Daerah penolakan**

  Gagal tolak H_0 karena $KS_{\text{hitung}} = 0,268 < KS_{\text{tabel}} = 0,338$

- **Kesimpulan** : Data sampel berdistribusi Binomial

Tabel Nilai Kritis Uji Kolmogorov-Smirnov

n	α = 0,20	α = 0,10	α = 0,05	α = 0,02	α = 0,01
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432
14	0,275	0,314	0,349	0,390	0,418
15	0,266	0,304	0,338	0,377	0,404
16	0,258	0,295	0,327	0,366	0,392
17	0,250	0,286	0,318	0,355	0,381

Uji distribusi Poisson (dengan KS)

Uji distribusi Poisson

$$Pi = P(x = i) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad , \lambda \approx \hat{\lambda} = rata - rata$$

- **Perumusan hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Poisson

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Poisson

Contoh 8

- Banyak pasien di ruang tunggu dalam kurun waktu atau interval per 30 detik. Uji apakah banyak pasien menunggu dalam interval per 30 detik mengikuti distribusi Poisson dengan parameter $\lambda = 3$?

Banyak pasien teramati	0	1	2	3	4	5	6	7	Total
Banyak interval	4	1	2	2	4	4	6	2	25

Jawab

- **Hipotesis**

  H_0 : data sampel berasal dari suatu populasi berdistribusi Poisson

  H_1 : data sampel tidak berasal dari suatu populasi berdistribusi Poisson

- **Statistik uji**

  r = banyaknya karakteristik yang diamati

  $$KS_{\text{hitung}} = \mathbf{max}\left|D\right| = 0,295$$

Poisson											
Banyak pasien teramati	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**		lambda
oi	4	1	2	2	4	4	6	2	**25**		3
e(-lambda)	0.049787	0.049787	0.049787	0.049787	0.049787	0.049787	0.049787	0.049787			
lambda^x	1	3	9	27	81	243	729	2187			
x!	1	1	2	6	24	120	720	5040			
pi	0.049787	0.149361	0.224042	0.224042	0.168031	0.100819	0.050409	0.021604	**0.988095**		
p observed	0.16	0.04	0.08	0.08	0.16	0.16	0.24	0.08			
F observed	0.16	0.2	0.28	0.36	0.52	0.68	0.92	1			
F expected	0.049787	0.199148	0.42319	0.647232	0.815263	0.916082	0.966491	0.988095			
|Fo-Fe|	0.110213	0.000852	0.14319	0.287232	**0.295263**	0.236082	0.046491	0.011905			
					KS Hitung						

- **Daerah penolakan**

  Tolak H_0 karena $KS_{\text{hitung}} = 0,295 > KS_{\text{tabel}} = 0,264$

- **Kesimpulan** : Data sampel tidak berdistribusi Binomial

n	α = 0,20	α = 0,10	α = 0,05	α = 0,02	α = 0,01
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
24	0,212	0,242	0,269	0,301	0,323
25	0,208	0,238	0,264	0,295	0,317
26	0,204	0,233	0,259	0,290	0,311

FAKULTAS TEKNOLOGI MAJU DAN MULTIDISIPLIN

Uji distribusi Normal (dengan Chi-Square, KS, dan Liliefors)

Uji Liliefors

Asumsi :

- sampel terdiri n pengamatan bebas
- skala pengukuran minimal yang mungkin digunakan nominal
- hasil pengamatan diklasifikasikan dalam r kategori yang tidak saling tumpang tindih

Uji Liliefors

- **Hipotesis:**

  H_0 : data sampel berasal dari distribusi normal

  H_1 : data sampel tidak berasal dari distribusi normal

- **Statistik uji**:

$$L_0 = \sup_{x} \left| F(z_i) - S(z_i) \right|$$

- **Daerah penolakan**:

  tolak Ho jika $L_0 > L_{\alpha, n}$

  $L_{\alpha, n}$ adalah nilai kritis untuk uji Liliefors

Uji Liliefors

Langkah-langkah:

1. Ubah x_i , i = 1, 2, ... , n ke dalam bentuk z_i , i = 1, 2, ... , n , melalui transformasi
2. Hitung $F(z_i) = P(z < z_i)$
3. Hitung proporsi $z_1 , z_2 , \ldots , z_n$ yang < z_i ; katakan $S(z_i)$ maka

$$S(z_i) = \frac{\text{banyaknya}\, z_1, z_2, \cdots, z_n \text{ yang} \leq z_i}{n}$$

4. Hitung $\left| F(z_i) - S(z_i) \right|$
5. Tentukan $L_0 = \sup_x \left| F(z_i) - S(z_i) \right|$
6. Bandingkan nilai L_0 dengan $L_{\alpha,n}$

Contoh 9

Berikut diberikan data :

23 27 33 40 48 48 57 59 62 68 69 70

yang diambil dari suatu populasi, akan diuji hipotesis nol bahwa sampel ini berasal dari populasi dengan distribusi normal pada $\alpha = 0.05$.

Jawab

Penyelesaian

PERUMUSAN HIPOTESIS :

H_0 : data sampel berasal dari distribusi normal

H_1 : data sampel tidak berasal dari distribusi normal

STATISTIK UJI : $L_0 = \underset{x}{Sup} \left| F(z_i) - S(z_i) \right|$

DAERAH KRITIS : tolak Ho jika $L_0 > L_{\alpha,\, n}$

Untuk $\alpha = 0.05$ dan $n = 12$ dari tabel nilai kritis uji Liliefors $L_{0.05\,,\,12} = 0{,}242$

Perhitungan :

Dari data di atas diperoleh : lihat table di slide selanjutnya

Jawab

i	1	2	3	4	5	6	7	8	9	10	11	12
data (urut)	23	27	33	40	48	48	57	59	62	68	69	70
mean												
50.33333												
std.dev												
16.54928												
z	-1.65163	-1.40993	-1.04738	-0.6244	-0.14099	-0.14099	0.402837	0.523688	0.704965	1.067519	1.127944	1.18837
F(z)	0.049305	0.07928	0.147463	0.266183	0.443938	0.443938	0.656466	0.699752	0.759584	0.857131	0.870328	0.882656
S(z)	0.083333	0.166667	0.25	0.333333	0.416667	0.5	0.583333	0.666667	0.75	0.833333	0.916667	1
|F(z)-S(z)|	0.034029	0.087387	0.102537	0.06715	0.027271	0.056062	0.073133	0.033086	0.009584	0.023798	0.046338	**0.117344**
												Lo

Dari tabel di atas tampak pada = 70 memberikan nilai   terbesar sehingga L_0 = 0,117.

Dari tabel nilai kritis uji Liliefors $L_{0,05\,,\;12}$ = 0,242 berarti $L_0 < L_{0,05\,,\;12}$ maka hipotesis nol diterima.

Kesimpulannya adalah bahwa populasi asal berdistribusi normal

Catatan :

Untuk pengujian keselarasan ini data harus dalam keadaan terurut dari kecil ke besar.

Tabel Nilai Kritis Untuk Uji Liliefors

Ukuran Sampel	Taraf Nyata (α)				
	0.01	0.05	0.10	0.15	0.20
n = 4	0.417	0.381	0.352	0.319	0.300
5	0.405	0.337	0.315	0.299	0.285
6	0.364	0.319	0.294	0.277	0.265
7	0.348	0.300	0.276	0.258	0.247
8	0.331	0.285	0.261	0.244	0.233
9	0.311	0.271	0.249	0.233	0.223
10	0.294	0.258	0.239	0.224	0.215
11	0.284	0.249	0.230	0.217	0.206
12	0.275	$\boxed{0.242}$	0.223	0.212	0.199
13	0.268	0.234	0.214	0.202	0.190
14	0.261	0.227	0.207	0.194	0.183
15	0.257	0.220	0.201	0.187	0.177
16	0.250	0.213	0.195	0.182	0.173
17	0.245	0.206	0.289	0.177	0.169
18	0.239	0.200	0.184	0.173	0.166
19	0.235	0.195	0.179	0.169	0.163
20	0.231	0.190	0.174	0.166	0.160
25	0.200	0.173	0.158	0.147	0.142
30	0.187	0.161	0.144	0.136	0.131
n > 30	$\frac{1.031}{\sqrt{n}}$	$\frac{0.886}{\sqrt{n}}$	$\frac{0.805}{\sqrt{n}}$	$\frac{0.768}{\sqrt{n}}$	$\frac{0.736}{\sqrt{n}}$

Sumber : Sudjana (1992)

Latihan Soal

Kerjakan Contoh 9 menggunakan Uji Chi-Square dan Uji KS

Terima Kasih