



Analisis Data: Test of Significance

Arrianto Mukti Wibowo

Sesi 9

Agenda

- Hypothesis testing

Remember: Inferential Statistics

- Sampel dipakai untuk menggeneralisasi, membuat kesimpulan mengenai populasi
- Ingat: bahwa sampel \neq populasi.
- Sampel \subset populasi
- Inductive thinking (vs. deductive thinking)

Statistical Testing Procedures

- State Null Hypothesis
- Choose the statistical test
 - How sample is drawn
 - How big is the sample
 - Measures of scale (nominal, ordinal, rank, ratio)
- Select desired level of significance
 - Umumnya 0.05 (5%)
- Compute the calculated difference value
- Obtain the critical test value
- Interpret the test



Pemahaman Dasar

Apakah masih rata-rata atau diluar kisaran rata-rata?

Null Hypothesis

- Kasus: Honda Insight dengan konsumsi bahan bakar 50 miles per gallon (mpg)
- Dari percobaan yang dilakukan oleh tim quality testing Honda terhadap 25 unit mobil Insight, diketahui betul bahwa konsumsinya 50 mpg.
- Sudin membeli mobil Honda Insight, tapi setelah mencatat pemakaian bahan bakarnya kok berbeda (misalnya ternyata 47 mpg).
- Pikir Sudin, “Hmm... ini sebenarnya masih masuk kisaran konsumsi BBM yang dijanjikan Honda apa tidak ya?”

H_0 dan H_A

- Maka Sudin membuat dua macam hipotesis
 - Null hypothesis (H_0): *“tidak ada perbedaan signifikan antara pengujian Honda (50 mpg) dengan catatan si Sudin.*
 - $H_0:\mu = 50$ mpg
 - Alternate hypothesis (H_A): *“ada perbedaan signifikan antara pengujian Honda (50 mpg) dengan catatan si Sudin.*
 - $H_A:\mu \neq 50$ mpg
 - Pakai two-tailed test
-
- Tapi, kalau ingin menguji apakah lebih rendah dari 50 mpg, agak berbeda pengujiannya.
 - $H_0:\mu = 50$ mpg
 - $H_A:\mu < 50$ mpg
 - Pake one-tailed test

Two-tailed test

- Asumsi mean populasi $\mu = 50$ mpg
- Standar deviasi populasi $\sigma = 10$ mpg
(kalau tidak bisa diambil dari sampel populasi, bisa ambil dari penelitian terdahulu, atau stdev dari sampel)
- Jumlah sampel $n = 25$ unit mobil
- Standar error of mean $\frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$
- Level of significance (α) adalah probabilitas menolak null sesungguhnya.
- Kalau diinginkan *confidence interval* 95%, (dimana *level of significancenya* $\alpha = 5\% = 0.05$) maka error di setiap ekor adalah $0.05/2 = 0.025 = 2,5\%$

Perhitungan *Critical Values*

- Ingin dicari critical value \bar{X}_c
- Jika significance level = 5% (0.05), maka Z-score *two-tailed* dapat dilihat pada tabel statistik, $Z = 1.96$ (plus dan negatif)

- Maka dapat dihitung dengan rumus $Z = \frac{\bar{X}_c - \mu}{\sigma_{\bar{X}}}$
- Hitung \bar{X}_c

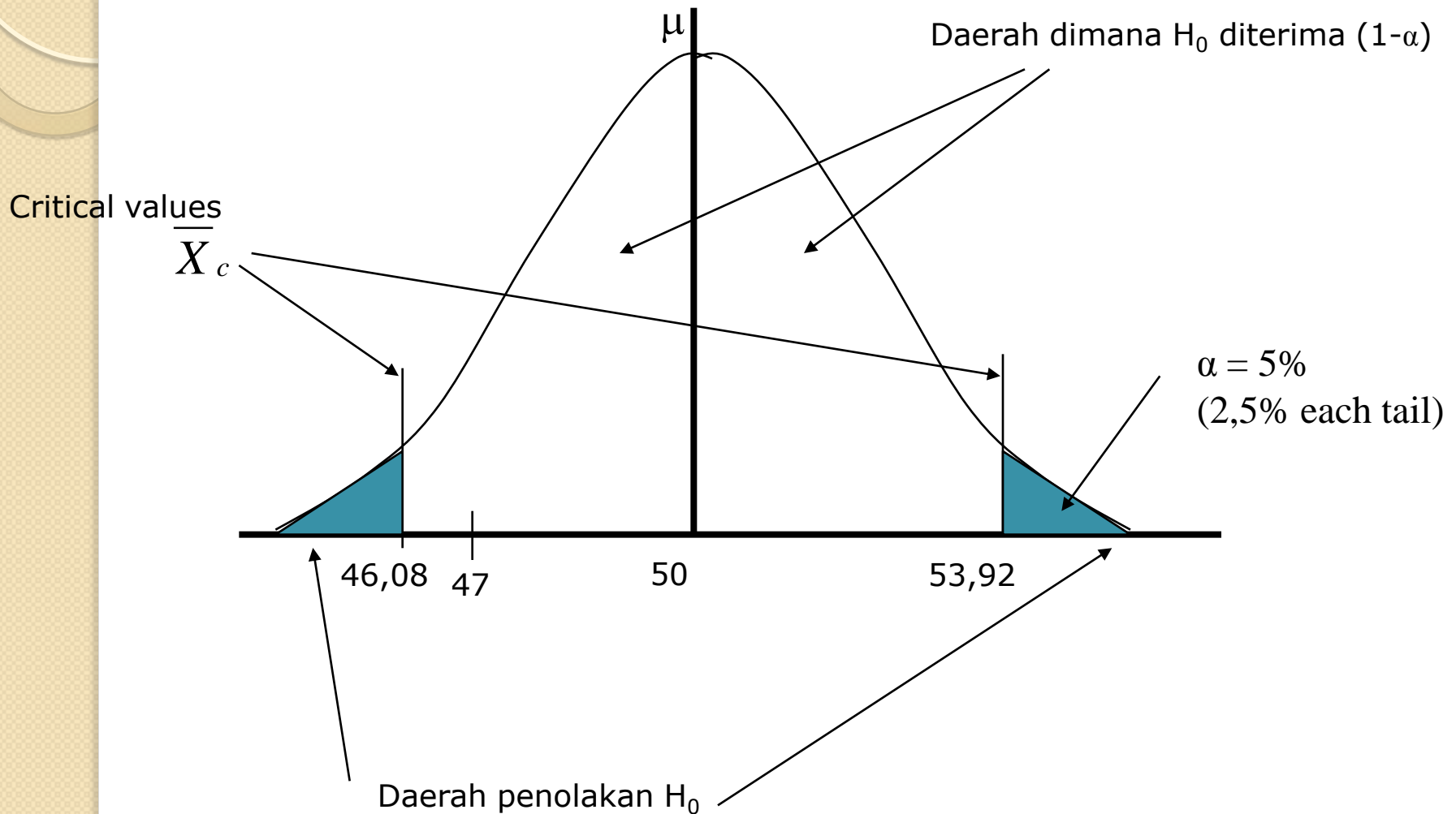
$$-1.96 = \frac{\bar{X}_{c-left} - 50}{2}$$

$$\bar{X}_{c-left} = 46.08$$

$$1.96 = \frac{\bar{X}_{c-right} - 50}{2}$$

$$\bar{X}_{c-right} = 53.92$$

Two-tailed test (sambungan)



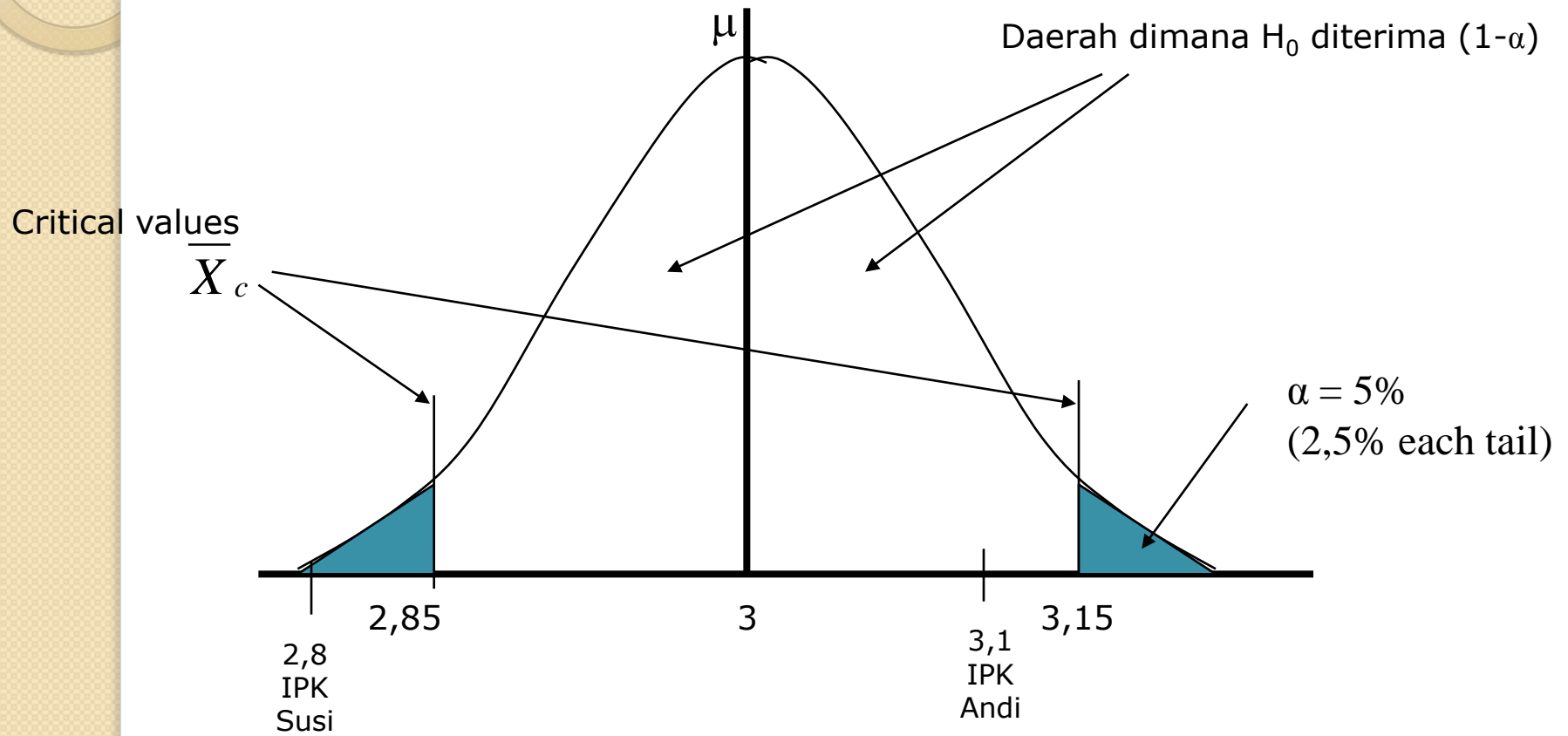
“Pemaknaan Honda Insight” si Sudin

- Jika mobil Insight si Sudin konsumsi BBM-nya 47 mpg, sebenarnya berarti bahwa Honda tidak menipu...!
- Tapi kalau ternyata mobil Insight si Sudin konsumsi BBM-nya 45 miles per gallon, nah... ketahuan bahwa Honda bo'ong 😞

Contoh pemaknaan lain “IPK di Fasilkom UI”

- Anda melakukan penelitian terhadap 40 orang mahasiswa dari populasi 800 mahasiswa Fasilkom UI. Ditemukan bahwa rata-rata IPK sampel adalah 3.
- IPK Andi ternyata 3,1. Hehehe... 😊 apakah Andi perlu girang bahwa IPK Anda lebih tinggi dari rata-rata? Belon tentu, jangan-jangan critical valuenya adalah 3,15 ($\alpha=5\%$). Dalam kasus seperti ini Andi sebenarnya cuma rata-rata! 😊
- Tapi kalo IPK Susi 2,8? Bisa jadi kalau critical value bawah-nya 2,85, yah... mohon maaf, Susi dibawah rata-rata 😞

IPK Andi & Susi



Kesimpulan

- Nah, setelah Anda melakukan penelitian nanti, Anda bisa tahu:
 - Apakah Anda itu makhluk *ajaib* atau *kodian*?
 - Ajaib kalo IPK-nya: 1,8 atau 3,8
 - Kodian kalo IPK-nya: 3,1 atau 3,0
 - Nge-kost itu ajaib atau kodian, yach?
 - Pake sendal ke kampus itu ajaib atau kodian, yach?



Type I & Type II Error

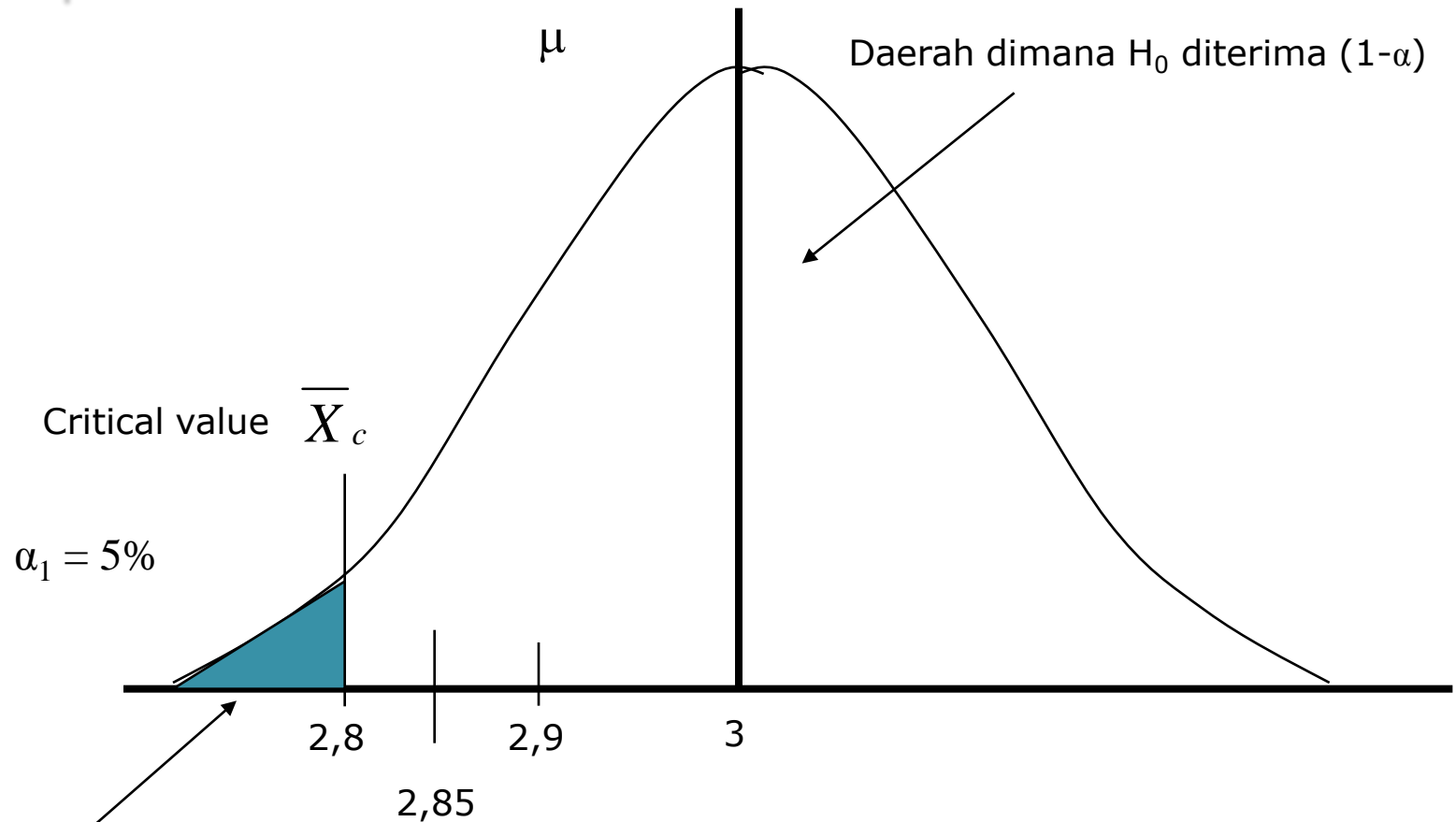
Type I Error

“Singkirkan orang yang tak disukai”

- Anda melakukan penelitian terhadap 40 orang mahasiswa dari populasi 800 mahasiswa Fasilkom UI. Ditemukan bahwa rata-rata IPK sampel adalah 3
- Misalnya untuk menjadi asisten dosen IPK-nya harus lebih besar dari rata-rata, dengan $\alpha = 5\%$, asumsikan ternyata *critical value* bawahnya adalah 2,8.
- Dalam kasus spesifik ini Anda harus pakai *one-tailed*
- Si Joni IPK-nya ternyata 2,85. Apakah dia sebenarnya bisa jadi asisten dosen? Ya.
- Hmm... tapi kalau dosen nggak suka dengan si Joni, bisa jadi si dosen memilih $\alpha = 10\%$, confidence interval 90%. Hasilnya: critical value bawahnya jadi 2,9.
- Apakah Joni diterima? Nggak...!
- Si Dosen melakukan kesalahan tipe I, bahwa sesungguhnya Joni harusnya diterima, tapi karena dosen α dibuat besar, maka Joni tidak diterima.
- Itulah kenapa pemilihan α tergantung dengan konteks. Kalau dalam pengujian obat α dibuat sekecil mungkin.

Type I Error on one-tailed

$$\alpha_1 = 5\%$$

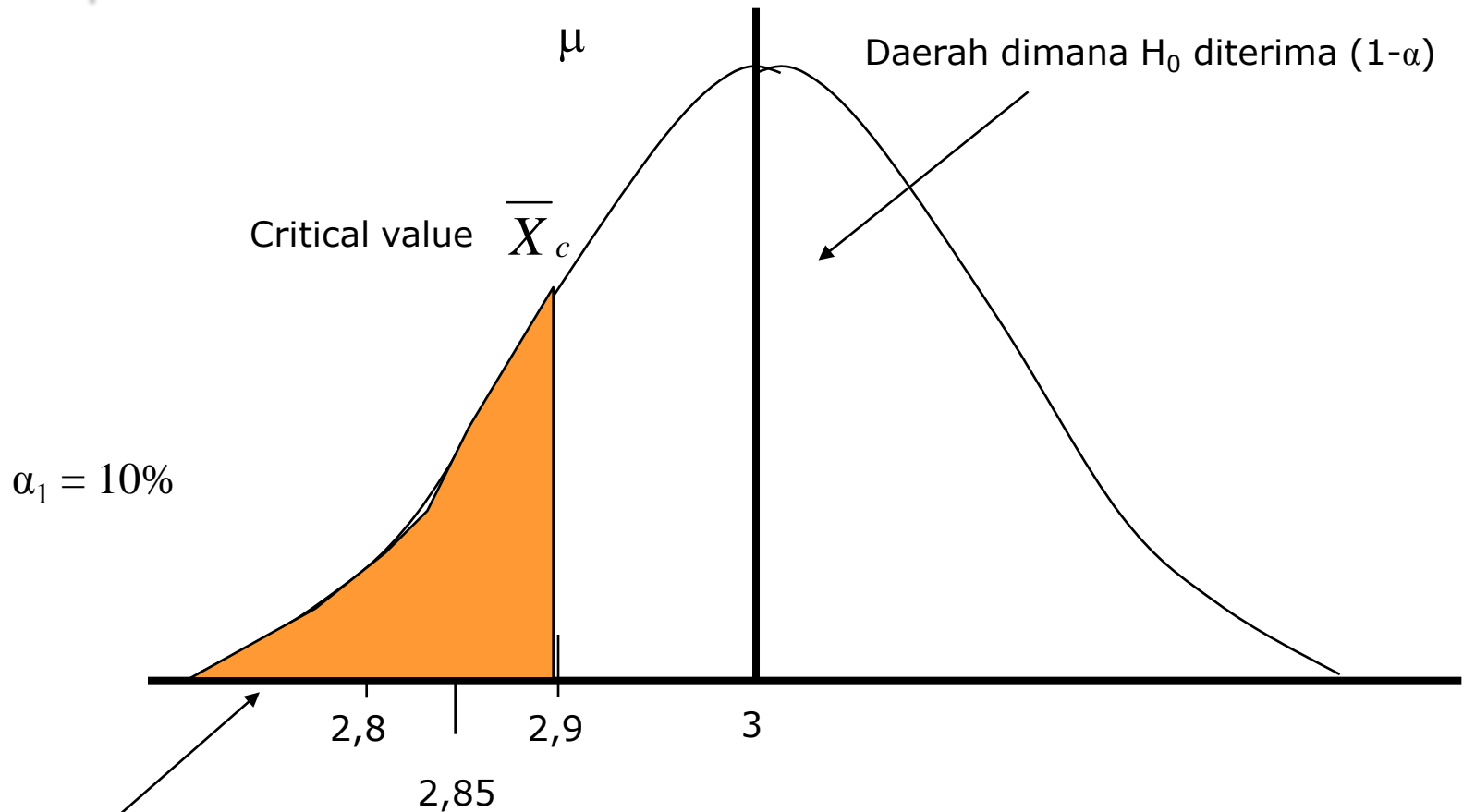


Joni diterima jadi asisten dosen

Daerah penolakan H_0

Type I Error on one-tailed

$$\alpha_1 = 10\%$$



Joni GAGAL jadi asisten dosen

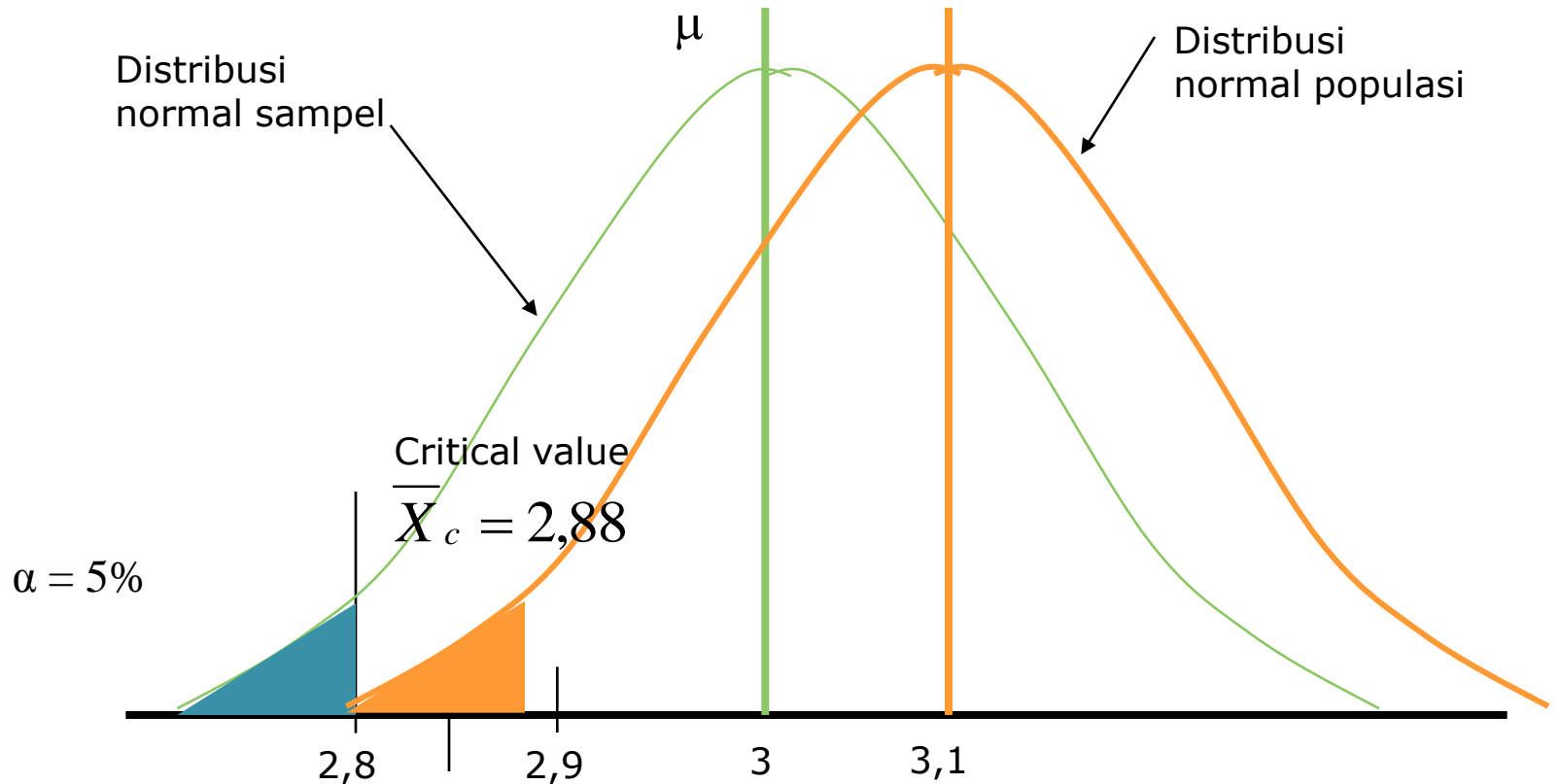
Daerah penolakan H_0

Type II error

- Terjadi manakala ternyata data mengenai mean sesungguhnya mengenai populasi diketahui.
- Misalnya pada setiap akhir tahun ajaran, Sub-Bagian Akademik (SBA) mengeluarkan statistik deskriptif yang berisi rata-rata IPK.
- Balik ke contoh Joni.
- Ternyata IPK dari sensus SBA rata-ratanya 3,1!
- Hmm... apa artinya buat Joni?
- Kita ketahui, untuk menjadi asisten dosen IPK-nya harus lebih besar dari rata-rata.

Type II Error on one-tailed

$$\alpha = 5\%$$



**Joni harusnya
tidak bisa jadi
asdos**

Joni kalau sudah diterima jadi asisten dosen, sesungguhnya si **dosen telah melakukan kesalahan**, karena ternyata IPK populasi bukan 3, tapi 3,1! Kenapa? Waktu mengangkat Joni, SBA belum mengeluarkan statistik populasi.



Aneka Macam Alat Uji Statistik (Statistical Testing)

Type of test

- Parametric test
- Non-parametric test

Asumsi untuk Parametric test

- Observasi harus independen: pengambilan satu kasus tidak berpengaruh pada kemungkinan kasus lain dimasukkan dalam sampel
- Observasi harus dilakukan pada sampel yang memiliki distribusi normal ($30 + \text{error } 10$)
- Harus menggunakan ukuran interval atau rasio, sehingga bisa dilakukan operasi aritmatika

Cara memilih pengujian

- Tergantung jenis skala: nominal, ordinal, interval/rasio
- Berapa sampel:
 - One sample: contoh kasus Honda Insight & IPK yang sudah dijelaskan sebelumnya
 - Two sample:
 - pengujian apakah ada perbedaan signifikan antara mahasiswa (sample pertama) dan mahasiswi (sample kedua) dalam keputusan untuk kost. Ini contoh pula untuk *independent sample*
 - *k*-sample: apakah ada perbedaan tingkat kepuasan user Linux, Windows dan Macintosh?
- Independent atau related sample:
 - Pengujian apakah ada perbedaan daya serap pembelajaran pada *orang yang sama* (related sample), dengan dan tanpa ngopi 'Tubruk'

Beberapa Alat Uji Statistik

Measures Scale	One-sample case	Two-test sample		k-test sample	
		Related sample	Independent sample	Related sample	Independent sample
Nominal	<input type="checkbox"/> Binomial <input type="checkbox"/> X ² one-sample test	<input type="checkbox"/> McNemar	<input type="checkbox"/> Fisher exact test <input type="checkbox"/> X ² two-sample test	<input type="checkbox"/> Cochran Q	<input type="checkbox"/> X ² for k-sample test
Ordinal	<input type="checkbox"/> Kolmogorov-Smirnov one-sample test <input type="checkbox"/> Runs test	<input type="checkbox"/> Sign test <input type="checkbox"/> Wilcoxon matched-pairs test	<input type="checkbox"/> Median test <input type="checkbox"/> Mann-Whitney U <input type="checkbox"/> Kolmogorov-Smirnov <input type="checkbox"/> Wald-Wolfowitz	<input type="checkbox"/> Freidman two-way ANOVA	<input type="checkbox"/> Median extension <input type="checkbox"/> Kruskal-Wallis one-way ANOVA
Interval & Ratio	<input type="checkbox"/> t-test <input type="checkbox"/> Z-test	<input type="checkbox"/> t-test for paired samples	<input type="checkbox"/> t-test <input type="checkbox"/> Z-test	<input type="checkbox"/> Repeated measures ANOVA	<input type="checkbox"/> One-way ANOVA <input type="checkbox"/> N-way ANOVA

Tujuannya

- Ingat, bahwa tujuan dari pengujian ini adalah untuk mengetahui ada perbedaan signifikan atau tidak!



Contoh:

Two independent sample test

No Smoking at BP!

- Misalnya perusahaan minyak “Bau Puntung” (BP) ingin mengetahui apakah merokok berdampak pada kecelakaan pegawainya.
- BP selalu mencatat kecelakaan saat bekerja, jadi dapatlah dibuat sebuah sampel dari orang yang pernah mengalami kecelakaan dan orang yang tidak pernah mengalami kecelakaan.
- Kemudian BP menanyakan setiap orang tersebut, apakah dia:
 - Perokok berat
 - Biasa (moderate)
 - Non-smoker

Tabel 2x3

On-the job accident

Smoker

Cell designation Count Expected value	Yes	No	Total
Heavy smoker	(1,1) 12 8,24	(1,2) 4 7,75	16
Moderate	(2,1) 9 7,73	(2,2) 6 7,27	15
Non-smoker	(3,1) 13 18,03	(3,2) 22 16,97	35
Total	34	32	66

Penjelasan

- Dalam teknik ini kita menguji perbedaan signifikan antara apa yang diobservasi vs apa yang diharapkan berdasarkan null-hypothesis.
- Kalau makin besar perbedaan antara yang diharapkan (*expected*) dengan yang diobservasi (*observed*), maka kemungkinannya memang bahwa perbedaan itu memang betul-betul berbeda.

Pengujian dengan χ^2

1. Null hypothesis (permasalahan disederhanakan):
 - H_0 : Tidak ada hubungan antara kecelakaan kerja dengan merokok atau tidaknya seorang pegawai
 - H_A : Ada hubungan antara kecelakaan kerja dengan merokok atau tidaknya seorang pegawai
2. Gunakan χ^2 .
3. Significance level $\alpha = 5\%$. *Degrees of freedom* (d.f.)
 $= (3-1)(2-1)=2$
4. Expected distribution dapat dilihat pada total pada margin tabel.
 - Kalau tidak ada hubungan antara kecelakaan kerja dan merokok, maka tentu akan ada proporsi perokok yang sama di dalam accident dan non-accident.
 - Expected value $(1,1) = 34 \times 16 / 66 = 8,24$

Tabel 2x3

On-the job accident

Smoker

Cell designation Count observed Expected value	Yes	No	Total
Heavy smoker	(1,1) 12 8,24	(1,2) 4 7,75	16
Moderate	(2,1) 9 7,73	(2,2) 6 7,27	15
Non-smoker	(3,1) 13 18,03	(3,2) 22 16,97	35
Total	34	32	66

Pengujian (lanjutan)

- Hitung $\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$

$$\chi^2 = \frac{(12 - 8,24)^2}{8,24} + \dots + \frac{(22 - 16,97)^2}{16,97} = 7,01$$

- Lihat dari tabel Chi-square, ternyata untuk $\alpha = 5\%$ dan $d.f.=2$, maka *critical value*-nya adalah 5,99.

- Karena kalkulasi $\chi^2 = 7,01 > 5,99$, maka H_0 ditolak. Artinya ada mungkin hubungan antara kecelakaan kerja dengan merokok.

H_0 diterima



Catatan

- Sesungguhnya yang dibuktikan disini adalah dalam tabel 2x3 itu ada perbedaan signifikan antara sel-sel-nya.
- Oleh karena itu, dari pada bingung, saya anjurkan nanti supaya lebih banyak menggunakan tabel 2x2.
- Tapi *sesungguhnya* tidak menjelaskan apakah ada hubungan antara *kebiasaan merokok dengan kemungkinan **lebih besar** mendapatkan kecelakaan kerja!*
- Lho, bagaimana?
- *Measures of association* akan dibahas nanti. Dalam kasus tabel 2x3 dengan skala nominal tadi, mungkin menggunakan *Cramer's V* atau *Contingency coefficient C*.
- Kalau 2x2, pergunakan phi (ϕ)
- Contoh *measures of association* yang lain adalah Pearson's *r* dalam liner regression.



Penutup Sesi

Penutupan Sesi

- Please download program statistik gratis, misalnya “Statistical Lab”
- Diskusi mengenai topik untuk kelompok yang melakukan survey
- Pengumuman mailing list yahoogroups