

Sistem Rekomendasi - A/B Testing

1. Dalam pengujian A/B yang dilakukan, Anda menggunakan rumus t-statistik. Apa alasan yang tepat dan mendasari pilihan ini dibandingkan menggunakan z-statistik?
 - a. Studi kasus bertujuan untuk membandingkan nilai rata-rata sehingga lebih cocok untuk t-statistik.
 - b. Studi kasus bertujuan membandingkan proporsi nilai dari masing-masing kelompok yang diuji.
 - c. Studi kasus bertujuan menghitung keseluruhan populasi yang mengikuti pola distribusi binomial.
 - d. Studi kasus bertujuan untuk menghitung suatu sampel dari seluruh populasi dengan menerapkan skema nilai acak berdistribusi Gaussian.
2. Dalam soal yang dikerjakan, Anda membuat fungsi `degrees_of_freedom` sebelum melakukan perhitungan t-statistic. Jelaskan mengapa tahapan menghitung derajat kebebasan perlu ditempuh terlebih dahulu?
 - a. Nilai degree of freedom menjadi pembatas untuk nilai t-statistik tidak melebihi nilai ini jika ingin *hipotesis null* diterima.
 - b. Dalam perhitungan derajat kebebasan melibatkan standar deviasi dan jumlah sampel untuk masing-masing kelompok pengujian. Kedua nilai tersebut digunakan untuk menghitung nilai t-statistic.
 - c. Degree of freedom memainkan peran kunci dalam memastikan dataset mendekati distribusi normal sebagai prasyarat untuk menghitung t-statistik dan z-statistik.
 - d. Soal tersebut memiliki kasus spesifik, yakni data yang digunakan mengikuti pola distribusi binomial sehingga degree of freedom digunakan untuk menekan nilai agar jatuh dalam interval $[0, 1]$
3. Manakah dari opsi di bawah ini yang merupakan rumus t-test yang digunakan dalam soal?

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

a.

$$t = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

b.

$$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

c.

$$t = 2 \times (1 - \text{CDF}(|t|))$$

d.

4. Dalam soal yang Anda kerjakan, hasil uji menunjukkan bahwa hipotesis nol ditolak. Apa interpretasi yang tepat berdasarkan konteks kasus tersebut?
 - a. Dalam uji hipotesis, disimpulkan bahwa terdapat perbedaan signifikan antara kedua kelompok. **Kelompok B yang menerima sistem rekomendasi menunjukkan hasil yang lebih tinggi** dibandingkan kelompok A dan perbedaan ini terbukti signifikan secara statistik serta tidak terjadi secara kebetulan. Jadi, sistem rekomendasi dapat diterapkan dalam semua versi aplikasi.
 - b. Uji hipotesis menunjukkan perbedaan signifikan antara kedua kelompok. Hal ini membuktikan **kelompok A yang memiliki nilai rata-rata lebih tinggi** dibanding kelompok B tidak terjadi secara kebetulan dan signifikan secara statistik. Jadi, sistem rekomendasi dapat diterapkan dalam semua versi aplikasi.
 - c. Uji hipotesis ditolak karena tingkat signifikansi 0,05 menunjukkan adanya perbedaan yang signifikan antara kedua kelompok. Ini berarti setiap hasil pengamatan memiliki peluang kesalahan sebesar $\pm 5\%$ dalam interpretasinya.
 - d. Uji hipotesis ditolak karena perbedaan rata-rata pada masing-masing kelompok tidak sebesar 50%. Jadi, sistem rekomendasi yang dikembangkan tidak cocok untuk diimplementasikan untuk semua versi aplikasi.
5. Dalam kasus yang Anda tangani, p-value merepresentasikan probabilitas yang diperoleh nilai t-statistic yang **sama ekstremnya atau lebih ekstrem** dibandingkan nilai yang diamati. Dengan alasan tersebut, rumus mana yang cocok untuk menghitung nilai p-value?
 - a. $p = 2 \times (1 - \text{PDF}(|t|))$
 - b. $p = \text{CDF}(t)$
 - c. $p = 2 \times (1 - \text{CDF}(|t|))$
 - d. $p = \text{PDF}(|t|)$