STATISTIK INDUSTRI 1

Agustina Eunike, ST., MT., MBA

DISTRIBUSI SAMPLING

PENGANTAR

Distribusi Sampling

Distribusi Sampling

- Distribusi peluang suatu statistik:
  - ✓Distribusi sampling rataan
  - ✓Distribusi sampling proporsi
  - ✓Distribusi sampling variansi

Distribusi Sampling

DISTRIBUSI SAMPLING: RATAAN

Distribusi Sampling: Rataan

- Variabel acak: rata-rata sampel $\bar{X}$
- Distribusi sampling rataan $\bar{X}$ dengan ukuran sample n:
  - Distribusi yang dihasilkan dari eksperimen yang dilakukan berulang-ulang (selalu dengan ukuran sampel n) dan memberikan banyak nilai $\bar{X}$
  - Menggambarkan sebaran rata-rata sampel seputar rata-rata populasi μ
- Jika Populasi berdistribusi normal (μ, σ^2), maka Sampel acak yang diambil akan berdistribusi normal sama dengan populasinya.
- Jika **distribusi Populasi tidak diketahui** tetapi nilai μ dan σ^2 diketahui, pada kondisi jumlah sample acak besar $(n \geq 30)$, maka distribusi sampling $\bar{X}$ tetap **mendekati normal dengan rata-rata μ** dan **variansi σ^2/n**.
- Untuk mengetahui peluang rataan $\bar{X}$ dari **distribusi sampling normal / mendekati normal** dapat digunakan:
  - **CENTRAL LIMIT THEOREM**

CENTRAL LIMIT THEOREM

The Central Limit Theorem

As the sample size n increases without limit, the shape of the distribution of the sample means taken with replacement from a population with mean μ and standard deviation σ will approach a normal distribution. As previously shown, this distribution will have a mean μ and a standard deviation $\sigma/\sqrt{n}$.

Central Limit Theorem: If $\bar{X}$ is the mean of a random sample of size n taken from a population with mean μ and finite variance σ^2, then the limiting form of the distribution of

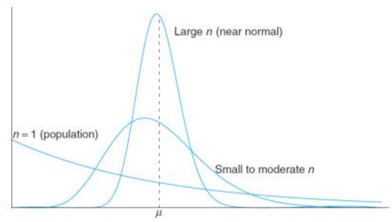
$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}},$$

as $n \to \infty$, is the standard normal distribution $n(z; 0, 1)$.

Distribusi Sampling: Rataan

- **CENTRAL LIMIT THEOREM**

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \qquad n \to \infty, distribusi\ normal\ standar\ n(z; 0,1)$$



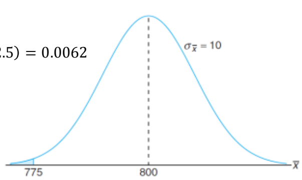
Distribusi Sampling: Rataan

- Contoh soal:
  - Diketahui usia produk lampu yang diproduksi oleh perusahaan ABC berdistribusi normal, dengan rata2 800 jam dan standard deviasi 40 jam. Hitung peluang 16 sampel yang diambil secara acak akan memiliki rata-rata usia produk kurang dari 775 jam.
  - Jawab:

$\mu_{\bar{X}} = 800; \sigma_{\bar{X}} = \frac{40}{\sqrt{16}} = 10$

$z = \frac{775-800}{10} = -2.5$

$P(\bar{X} < 775) = P(Z < -2.5) = 0.0062$



Distribusi Sampling: Rataan Dua Populasi

- Jika diketahui dua populasi dengan masing-masing rata-rata dan variansi adalah μ_1, σ_1^2 dan μ_2, σ_2^2. $\bar{X}_1$ adalah rata-rata sampel acak populasi pertama dengan ukuran sample n_1, dan $\bar{X}_2$ adalah rata-rata sampel acak populasi pertama dengan ukuran sample n_2. Jika kedua sampel acak tersebut independen, dan syarat pendekatan normal dipenuhi, maka perhitungan perbandingan peluang dua populasi $(\mu_1 - \mu_2)$, dapat dihitung dengan:

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 \text{ dan } \sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)}}$$

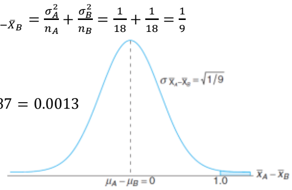
Distribusi Sampling: Rataan Dua Populasi

- Contoh soal:
  - Dua eksperimen dilakukan secara independen untuk membandingkan waktu kering dua jenis cat. Delapan plat dicat menggunakan cat tipe A, dan waktu kering yang diperlukan adalah 1 jam untuk masing-masing plat. Hal yang sama dilakukan pada cat tipe B. Standard deviasi kedua populasi diketahui sebesat 1 jam. Jika diasumsikan bahwa waktu kering kedua populasi adalah sama, hitung $P(\bar{X}_A - \bar{X}_B > 1)$, dengan $n_A = n_B = 18$
  - Jawab:

$\mu_{\bar{X}_A - \bar{X}_B} = \mu_A - \mu_B = 0$ dan $\sigma^2_{\bar{X}_A - \bar{X}_B} = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = \frac{1}{18} + \frac{1}{18} = \frac{1}{9}$

$z = \frac{1-(\mu_A - \mu_B)}{\sqrt{1/9}} = \frac{1-0}{\sqrt{1/9}} = 3.0$

$P(Z > 3.0) = 1 - P(Z < 3.0)$
$= 1 - 0.9987 = 0.0013$



Distribusi Sampling: Rataan Dua Populasi

- Latihan soal:
  - Jika diketahui informasi mengenai *lifetime* dua merk tabung telivisi merk A dan B sebagai berikut:

Population 1	Population 2
$\mu_1 = 6.5$	$\mu_2 = 6.0$
$\sigma_1 = 0.9$	$\sigma_2 = 0.8$
$n_1 = 36$	$n_2 = 49$

Berapakah peluang rata-rata *lifetime* sampling merk A paling sedikit 1 tahun lebih lama dibanding sampling merk B?

Distribusi Sampling

DISTRIBUSI SAMPLING: PROPORSI

Distribusi Sampling: Proporsi

- Variabel acak (p): proporsi kejadian sukses (x) dibanding total percobaan (n)

$$p = \frac{jumlah\ kejadian\ sukses}{jumlah\ percobaan} = \frac{x}{n}$$

  - Jika $n(1-\pi) \geq 5$, maka pendekatan dengan distribusi normal dapata dilakukan.
    - $\pi = P = proporsi\ populasi$
    - $n = ukuran\ sampel$

- **Sampling distribution of the proportion, p:**

  Mean = $E(p) = \pi$

  Standard error = $\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$ where π = population proportion, n = sample size

- **z-score for a given value of p:**

  $z = \frac{p-\pi}{\sigma_p}$ where p = the sample proportion value of interest

Distribusi Sampling: Proporsi

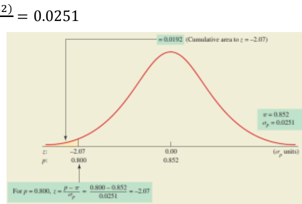
- Contoh soal:
  - Berdasarkan hasil sensus diketahui bahwa 85.2% penduduk dewasa di Kota Malang berpendidikan minimal SMA. Berapakah peluang tidak lebih dari 80% dari 200 penduduk dewasa kota Malang yang dipilih secara acak berpendidikan minimal SMA?
  - Jawab:

$\pi = P = 0.852; \quad p = 0.8; \quad n = 200$

$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0.852(1-0.852)}{200}} = 0.0251$

$z = \frac{p-\pi}{\sigma_p} = \frac{0.8-0.852}{0.0251} = -2.07$

$P(p \leq 0.8) = P(Z \leq -2.07)$
$= 0.0192$



Distribusi Sampling: Proporsi Dua Populasi

- Pada distribusi sampling beda dua proporsi berlaku hal-hal sbb:
  - Rata-rata: $\mu_{p1-p2} = P_1 - P_2$
  - Std deviasi: $\sigma_{p1-p2} = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$
  - Jika n_1 dan n_2 ($n_1, n_2 \geq 30$) cukup besar, distribusi sampling proporsi akan mendekati distribusi normal, dengan variabel random standar yang rumus Z-nya:

$$Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sigma_{p1-p2}}$$

Distribusi Sampling: Rataan Dua Populasi

- Contoh soal:
  - Berdasarkan sebuah penelitian, 1% orang yang tidak merokok terkena TBC dan dari populasi orang perokok, 5% orang di antaranya terkena TBC. Jika diambil sampel masing-masing 100 orang dari populasi orang merokok dan populasi orang tidak merokok yang terkena TBC lebih besar dari 5%?

Distribusi Sampling: Rataan Dua Populasi

- P1 = proporsi populasi perokok yang terkena TBC
- P2 = proporsi populasi bukan perokok yang terkena TBC

$\mu_{p1-p2} = P_1 - P_2$ = 5% - 1% = 4%

$$\sigma_{p1-p2} = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}} = \sqrt{\frac{0{,}05(1-0{,}05)}{100} + \frac{0{,}01(1-0{,}01)}{100}}$$

$$Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sigma_{p1-p2}} = \frac{5\% - 4\%}{2{,}4\%} = 0{,}42$$

Distribusi Sampling: *Finite Population*

- Jika sampling dilakukan tanpa pergantian dan pada populasi yang finite, maka perhitungan standard deviasi akan berbeda, menjadi:

- **Standard error for the sample mean when sampling without replacement from a finite population:**

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

where n = sample size
N = population size
σ = population standard deviation

- **Standard error for the sample proportion when sampling without replacement from a finite population:**

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

where n = sample size
N = population size
π = population proportion

Distribusi Sampling

DISTRIBUSI SAMPLING: VARIANSI

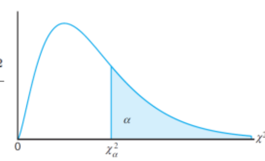
Distribusi Sampling: Variansi

- Variabel acak: variansi sampel S^2
- Distribusi sampling variansi S^2 dengan ukuran sample n:
  - ➢Distribusi yang dihasilkan dari eksperimen yang dilakukan berulang-ulang (selalu dengan ukuran sampel n) dan memberikan banyak nilai S^2
  - ➢Memberikan informasi sebaran nilai variansi s^2 sebagai inferensi nilai σ^2
  - ➢S^2 dapat dihitung dengan menggunakan rumus:

$$S^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right]$$

Distribusi Sampling: Variansi

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2}$$



chi-squared distribution with $v = n - 1$ degrees of freedom

Distribusi Sampling: Variansi

A manufacturer of car batteries guarantees that the batteries will last, on average, 3 years with a standard deviation of 1 year. If five of these batteries have lifetimes of 1.9, 2.4, 3.0, 3.5, and 4.2 years, should the manufacturer still be convinced that the batteries have a standard deviation of 1 year? Assume that the battery lifetime follows a normal distribution.

$$s^2 = \frac{(5)(48.26) - (15)^2}{(5)(4)} = 0.815.$$

$$\chi^2 = \frac{(4)(0.815)}{1} = 3.26$$

is a value from a chi-squared distribution with 4 degrees of freedom. Since 95% of the χ^2 values with 4 degrees of freedom fall between 0.484 and 11.143, the computed value with $\sigma^2 = 1$ is reasonable, and therefore the manufacturer has no reason to suspect that the standard deviation is other than 1 year. ∎

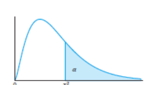


Table A.5 Critical Values of the Chi-Squared Distribution

	α									
v	**0.995**	**0.99**	**0.98**	**0.975**	**0.95**	**0.90**	**0.80**	**0.75**	**0.70**	**0.50**
1	0.0^4393	0.0^3157	0.0^3628	0.0^3982	0.00393	0.0158	0.0642	0.102	0.148	0.455
2	0.0100	0.0201	0.0404	0.0506	0.103	0.211	0.446	0.575	0.713	1.386
3	0.0717	0.115	0.185	0.216	0.352	0.584	1.005	1.213	1.424	2.366
4	0.207	0.297	0.429	0.484	0.711	1.064	1.649	1.923	2.195	3.357
5	0.412	0.554	0.752	0.831	1.145	1.610	2.343	2.675	3.000	4.351
6	0.676	0.872	1.134	1.237	1.635	2.204	3.070	3.455	3.828	5.348
7	0.989	1.239	1.564	1.690	2.167	2.833	3.822	4.255	4.671	6.346
8	1.344	1.647	2.032	2.180	2.733	3.490	4.594	5.071	5.527	7.344
9	1.735	2.088	2.532	2.700	3.325	4.168	5.380	5.899	6.393	8.343
10	2.156	2.558	3.059	3.247	3.940	4.865	6.179	6.737	7.267	9.342
11	2.603	3.053	3.609	3.816	4.575	5.578	6.989	7.584	8.148	10.341
12	3.074	3.571	4.178	4.404	5.226	6.304	7.807	8.438	9.034	11.340
13	3.565	4.107	4.765	5.009	5.892	7.041	8.634	9.299	9.926	12.340
14	4.075	4.660	5.368	5.629	6.571	7.790	9.467	10.165	10.821	13.339
15	4.601	5.229	5.985	6.262	7.261	8.547	10.307	11.037	11.721	14.339
16	5.142	5.812	6.614	6.908	7.962	9.312	11.152	11.912	12.624	15.338
17	5.697	6.408	7.255	7.564	8.672	10.085	12.002	12.792	13.531	16.338
18	6.265	7.015	7.906	8.231	9.390	10.865	12.857	13.675	14.440	17.338
19	6.844	7.633	8.567	8.907	10.117	11.651	13.716	14.562	15.352	18.338
20	7.434	8.260	9.237	9.591	10.851	12.443	14.578	15.452	16.266	19.337

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
6	7.231	7.841	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457
7	8.383	9.037	9.803	12.017	14.067	16.013	16.622	18.475	20.278	24.321
8	9.524	10.219	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.124
9	10.656	11.389	12.242	14.684	16.919	19.023	19.679	21.666	23.589	27.877
10	11.781	12.549	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588
11	12.899	13.701	14.631	17.275	19.675	21.920	22.618	24.725	26.757	31.264
12	14.011	14.845	15.812	18.549	21.026	23.337	24.054	26.217	28.300	32.909
13	15.119	15.984	16.985	19.812	22.362	24.736	25.471	27.688	29.819	34.527
14	16.222	17.117	18.151	21.064	23.685	26.119	26.873	29.141	31.319	36.124
15	17.322	18.245	19.311	22.307	24.996	27.488	28.259	30.578	32.801	37.698
16	18.418	19.369	20.465	23.542	26.296	28.845	29.633	32.000	34.267	39.252
17	19.511	20.489	21.615	24.769	27.587	30.191	30.995	33.409	35.718	40.791
18	20.601	21.605	22.760	25.989	28.869	31.526	32.346	34.805	37.156	42.312
19	21.689	22.718	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.819
20	22.775	23.828	25.038	28.412	31.410	34.170	35.020	37.566	39.997	45.314

Referensi

- Walpole, Ronald B., Myers, Raymond H., Myers, Sharon L., Ye, Keying, ***Probability & Statistics for Engineers and Scientist***, 9th ed, Prentice Hall Int., New Jersey, 2012.
- Weiers, R.M., 2011, ***Introduction to Business Statistics,*** Cengage Learning, OH, 2008.