Interval Estimation (contd..)

Summary:

$F_X(x; \theta, \theta_1, \theta_2, \ldots, \theta_k)$

- Let X be population r.v. with an unknown population parameter θ in its d.f. Let $(x_1, x_2, \ldots, x_n)$ be a random sample of size n. To find an interval estimate of θ.

- (a, b) where $a = a(x_1, x_2, \ldots, x_n)$ and $b = b(x_1, x_2, \ldots, x_n)$ is called an $100 \times (1-\epsilon)\%$ confidence interval of θ

$$P(A < \theta < B) = 1 - \epsilon \quad \forall\ \theta \in \Theta \text{ (set of admissible values of } \theta)$$

$$0 < \epsilon < 1$$

A, B are the random variables corresponding to a, b respectively.

- $\epsilon = 0.05$, $1-\epsilon = 0.95$, 95% - confidence interval

If we take 100 random samples of size n then we get 100 intervals (a_i, b_i), $i = 1, \ldots, 100$

Out of these 100 intervals approx. 95 intervals will contain θ.

- $X \sim N(m, \sigma)$. To find an interval estimate of m.

<u>Case I</u>: σ is known.

$(x_1, x_2, \ldots, x_n)$: random sample of size.

Choose statistic: $u = \dfrac{\bar{x} - m}{\sigma/\sqrt{n}}$

whose sampling distribution is $N(0,1)$.

To find $100 \times (1-\epsilon)\%$ confidence interval, we find two point $\pm u_\epsilon$ (from the Table of Standard Normal distri) s.t.

$$P\left(-u_\epsilon < U < u_\epsilon\right) = 1-\boxed{\epsilon}$$

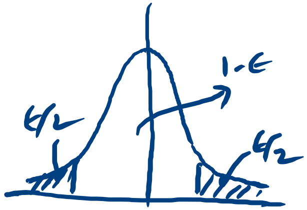
$$\Rightarrow P\left(-u_\epsilon < \frac{\bar{X}-m}{\sigma/\sqrt{n}} < u_\epsilon\right) = 1-\epsilon$$

$$\Rightarrow P\left(\bar{X}-\frac{\sigma u_\epsilon}{\sqrt{n}} < m < \bar{X}+\frac{\sigma u_\epsilon}{\sqrt{n}}\right) = 1-\epsilon$$

$100 \times (1-\epsilon)\%$ C.I. : $\left(\bar{x}-\frac{\sigma u_\epsilon}{\sqrt{n}}, \bar{x}+\frac{\sigma u_\epsilon}{\sqrt{n}}\right)$

Case II : σ unknown

We choose the statistic : (1) $t = \dfrac{\bar{x} - m}{s/\sqrt{n}}$



where $s^2 = \dfrac{n}{n-1} S^2$

Sampling distribution of t is t-distribution with $(n-1)$ degrees of freedom

To find $100 \times (1-\epsilon)\%$ confidence interval for m, we find two real nos. $\pm t_\epsilon$ (from the table of T-distribution) s.t.

$$P\left(-t_\epsilon < \frac{\bar{x} - m}{s/\sqrt{n}} < t_\epsilon\right) = 1 - \epsilon$$

$$\Rightarrow P\left(\bar{x} - \frac{s\, t_\epsilon}{\sqrt{n}} < m < \bar{x} + \frac{s\, t_\epsilon}{\sqrt{n}}\right) = 1-\epsilon$$

$100\times(1-\epsilon)\%$ C.I. : $\left(\bar{x} - \frac{s\, t_\epsilon}{\sqrt{n}},\ \bar{x} + \frac{s\, t_\epsilon}{\sqrt{n}}\right)$

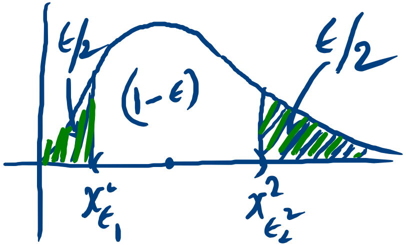
Prob: If your sample is : $\boxed{160, 161, 150, 170, 172}$

$X \sim N(m, \sigma)$ unknown

choose: $t = \dfrac{\bar{x} - m}{\boxed{s}/\sqrt{n}} \sim$?

Sampling distribution of t is t-distribution with $5-1=4$ degrees of freedom

② Find $100\times(1-\epsilon)\%$ confidence interval for the parameter σ of a normal (m,σ) population:



Sol. $X \sim$ normal (m,σ)

unknown parameters

$(x_1, x_2, \ldots, x_n)$: random sample of size n

Choose the statistic $\chi^2 = \dfrac{nS^2}{\sigma^2}$

whose sampling distribution is $\chi^2(n-1)$

Choose two real no.s $\chi^2_{\epsilon_1}$ & $\chi^2_{\epsilon_2}$ from the Table of χ^2-distribution s.t.

$$P\left(\chi^2_{\epsilon_1} < \chi^2 < \chi^2_{\epsilon_2}\right) = 1-\epsilon$$

$$\Rightarrow P\left(\chi^2_{\epsilon_1} < \frac{nS^2}{\sigma^2} < \chi^2_{\epsilon_2}\right) = 1-\epsilon$$

$$\Rightarrow P\left(\frac{1}{\chi^2_{\epsilon_1}} > \frac{\sigma^2}{nS^2} > \frac{1}{\chi^2_{\epsilon_2}}\right) = 1-\epsilon$$

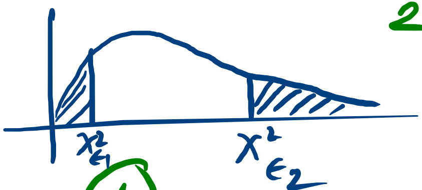
$$\Rightarrow P\left(S\sqrt{\frac{n}{\chi^2_{\epsilon_2}}} < \sigma < S\sqrt{\frac{n}{\chi^2_{\epsilon_1}}}\right) = 1-\epsilon$$

$(1-\epsilon)100\%$ C.I : $\left(S\sqrt{\frac{n}{\chi^2_{\epsilon_2}}}\ ,\ S\sqrt{\frac{n}{\chi^2_{\epsilon_1}}}\right)$.

To choo $\chi^2_{\epsilon_1}$ and $\chi^2_{\epsilon_2}$ we consider:

$$P(0<\chi^2<\chi^2_{\epsilon_1}) = P(\chi^2_{\epsilon_2}<\chi<\infty)$$

$$= \frac{P(0<\chi^2<\chi^2_{\epsilon_1}) + P(\chi^2_{\epsilon_2}<\chi<\infty)}{2} = \frac{\epsilon}{2}$$



$$\Rightarrow \quad P(0<\chi^2<\chi^2_{\epsilon_1}) = \frac{\epsilon}{2}$$

$$P(\chi^2>\chi^2_{\epsilon_1}) = 1-\frac{\epsilon}{2} \quad \text{——— } (1)$$

and $$P(\chi^2>\chi^2_{\epsilon_2}) = \frac{\epsilon}{2}. \quad \text{——— } (2)$$

③ Find an $\boxed{\text{approximate}}$ $100 \times (1-\epsilon)\%$ confidence interval for p for a binomial (N, p) population.

$X \sim$ binomial (N, p)

Consider a random sample of size 1 : x_1

From De Moire Laplace Limit Th,

$$U = \frac{X - Np}{\sqrt{Npq}} \quad \text{is approximately normal } (0,1).$$

A good estimate (MLE) of p is $\hat{p} = \frac{\bar{X}}{N} = \frac{x_1}{N}$

Thus $\sqrt{Npq} = \sqrt{N\frac{x_1}{N}\left(1-\frac{x_1}{N}\right)} = \sqrt{x_1(N-x_1)/N}$

Thus $\dfrac{X-Np}{\sqrt{x_1(N-x_1)/N}}$ is approximately $N(0,1)$

For finding $100\times(1-\epsilon)\%$ C.I.

We choose points $\pm u_\epsilon$ from the table of $N(0,1)$

s.t.

$$P\left(-u_\epsilon < \frac{X-Np}{\sqrt{x_1(N-x_1)/N}} < u_\epsilon\right) \simeq 1-\epsilon .$$

$$\Rightarrow P\left(\frac{X}{N} - u_\epsilon\sqrt{\frac{x_1(N-x_1)}{N^3}} < p < \frac{X}{n} + u_\epsilon\sqrt{\frac{x_1(N-x_1)}{N^3}}\right) \simeq 1-\epsilon$$

$100\times(1-\epsilon)\%$ C.I. $\left(\frac{x_1}{N} - u_\epsilon\sqrt{\frac{x_1(N-x_1)}{N^3}}, \frac{x_1}{N} + u_\epsilon\sqrt{\frac{x_1(N-x_1)}{N^3}}\right)$

④ Find 95% confidence Interval for the mean of a normal distribution $\boxed{\text{with } \sigma = 3}$, given the random sample of size 4 : $(2.3, -0.2, -0.4, -0.9)$.

[Given $P(U > 1.96) = 0.025$ where $U \sim N(0,1)$]

Sol. Use the method describe before choosing $u = \dfrac{\bar{x} - m}{\sigma/\sqrt{n}}$.

95% C.I. is $\left(\bar{x} - \dfrac{\sigma u_\epsilon}{\sqrt{n}}, \; \bar{x} + \dfrac{\sigma u_\epsilon}{\sqrt{n}} \right)$

$\epsilon = 0.05$, $\epsilon/2 = 0.025$, $u_\epsilon = 1.96$, $n = 4$,

$\bar{x} = \dfrac{2.3 - 0.2 - 0.4 - 0.9}{4} =$

(8)

x_i	$y_i = x_i - 68$	y_i^2
63	-5	25
66	-2	4
63	-5	25
67	-1	1
68	0	0
69	1	1
70	2	4
71	3	9
72	4	16
71	3	9

$\bar{x}$, S^2

Linear Tran

$x_i = a y_i + b$

$\boxed{a = 1 , b = -68}$

$\bar{x} = a \bar{y} + b$

$$S_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= a^2 S_y^2$$

$$\bar{x} = \frac{1}{10} \sum^{10} y_i + 68 = 0 \cancel{+}68 = 0$$

$$S_x^2 = \frac{1}{10} \sum_{i=1} y_i^2 - (\bar{y})^2 = \frac{94}{10} - 0 = 9.4$$

For 95% C.I

$\epsilon = 0.05$

$\epsilon/2 = 0.025$

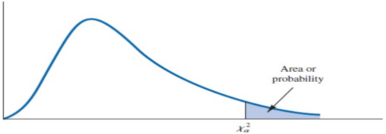
$1 - \epsilon/2 = 0.975$

If $n = 10$

$\chi^2_{\epsilon_1} = 2.700$

$\chi^2_{\epsilon_2} = 19.023$

TABLE 11.1 SELECTED VALUES FROM THE CHI-SQUARE DISTRIBUTION TABLE*



Degrees of Freedom	Area in Upper Tail .99	.975	.95	.90	.10	.05	.025	.01
1	.000	.001	.004	.016	2.706	3.841	5.024	6.635
2	.020	.051	.103	.211	4.605	5.991	7.378	9.210
3	.115	.216	.352	.584	6.251	7.815	9.348	11.345
4	.297	.484	.711	1.064	7.779	9.488	11.143	13.277
5	.554	.831	1.145	1.610	9.236	11.070	12.832	15.086
6	.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.878	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892
40	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691
60	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379
80	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329
100	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807

*Note: A more extensive table is provided as Table 3 of Appendix B.

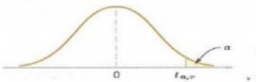


TABLE • V Percentage Points $t_{\alpha,\nu}$ of the t Distribution

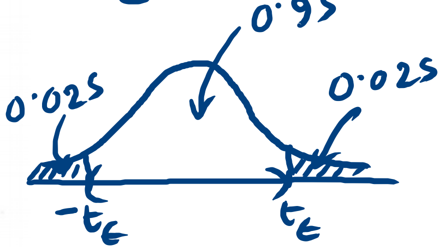
ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = degrees of freedom.

$t_{0.05,10} = 1.812$

95%

$\epsilon = 0.05$



$n = 20$

$t_{0.025,19} = 2.093$

✓ $t_\epsilon = 2.093$

✓ $-t_\epsilon = -2.093$