

Summary:

pdf. of  $\hat{\theta}^{\text{MLE}}$ .

Fisher information

$$I(\theta) = -\mathbb{E} \left[ \frac{\partial^2}{\partial \theta^2} (\log f(x; \theta)) \right]$$

asymptotically normal.

Prop:

$$\hat{\theta}^{\text{MLE}} \xrightarrow{d.} N(\theta, (nI(\theta))^{-1})$$

Confidence interval for population mean.

$$X_1, X_2, \dots, X_n. \quad \begin{array}{c} \mathbb{E}X \\ \parallel \\ n \end{array} \quad \begin{array}{c} \text{Var}(X) \\ \parallel \\ \sigma^2 \end{array}$$

To do interval estimation, one need to know the distribution.

①  $\sigma^2$  known,  $n$  is large

$$\frac{\bar{X}_n - \mathbb{E}X}{\sigma/\sqrt{n}} \xrightarrow{d.} N(0, 1)$$

$$P(\hat{\mu}_1 \leq \bar{X}_n \leq \hat{\mu}_2) = 0.95$$
$$\downarrow \quad \downarrow$$
$$\bar{X}_n - \quad \quad \quad \bar{X}_n + \quad .$$

rmk1. proof is via Central Limit Theorem.

rmk2. typically  $n \geq 25$  is sufficiently large to apply the approximation.

②  $\sigma^2$  unknown,  $X$  is normal

Sample variance  $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$   
 $\hat{\sigma}$  is called sample s.d.

$$\frac{\bar{X}_n - \mathbb{E}X}{\hat{\sigma}/\sqrt{n}} \sim t_{n-1}$$

rmk3. It is used for small  $n$  typically ( $n < 25$ ). The main reason is that

Prop:  $t_n \xrightarrow{d.} N(0, 1)$

Thus, when  $n$  is large, using normal table does not lead to large difference. Apply ③ would be better in the sense that " $X$  is normal" is not required.

rmk4. proof is via Basu's Theorem, which tells why  $\bar{X}_n$  and  $\hat{\sigma}^2$  are independent.

③  $\boxed{\sigma^2 \text{ unknown. } n \text{ large.}}$

$$\frac{\bar{X}_n - \mathbb{E}X}{\hat{\sigma}/\sqrt{n}} \xrightarrow{d.} N(0, 1)$$

$\hat{\sigma}$  can be any consistent estimator  
of  $\sigma$ . i.e.  
 $\hat{\sigma} \xrightarrow{P.} \sigma$

rmk 5. proof is via Slutsky's Theorem.

$$\text{CLT} \Rightarrow \left. \begin{array}{l} \frac{\bar{X}_n - \mathbb{E}X}{\sigma/\sqrt{n}} \xrightarrow{d.} N(0, 1) \\ \frac{\hat{\sigma}}{\sigma} \xrightarrow{P.} 1 \end{array} \right\} \xrightarrow{\text{Slutsky's}} \frac{\bar{X}_n - \mathbb{E}X}{\hat{\sigma}/\sqrt{n}} \xrightarrow{d.} N(0, 1)$$

rmk 6. From the sketch of the proof, one can see that "X is normal" is not required.

distribution of  $\bar{X}_n$ .

$n$ large	$X$ normal	$\sigma^2$ known	Solution
Yes	Yes	Yes	Standardisation
Yes	Yes	No	$t$ distribution (and $t_n \xrightarrow{d.} N(0, 1)$ )
Yes	No	Yes	CLT
Yes	No	No	CLT + consistent estimator of $\sigma^2$ Slutsky's Theorem.
No	Yes	Yes	Standardisation
No	Yes	No	$t$ distribution
No	No	Yes	{ Nothing can be done using knowledge from ST102 }
No	No	No	

Standardisation:  $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$

$t$  distribution:  $\frac{\bar{X}_n - \mu}{\hat{\sigma}/\sqrt{n}} \sim t_{n-1}$

CLT:  $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d.} N(0, 1) \text{ as } n \rightarrow \infty$

CLT + Consistent estimator of  $\sigma^2$ : 
$$\left. \begin{array}{l} \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d.} N(0, 1) \\ \frac{\hat{\sigma}}{\sigma} \xrightarrow{P.} 1 \end{array} \right\} \begin{array}{l} \text{Slutsky's} \\ \text{Theorem} \end{array} \Rightarrow \frac{\bar{X}_n - \mu}{\hat{\sigma}/\sqrt{n}} \xrightarrow{d.} N(0, 1)$$

How to construct consistent estimators ?

Law of Large Numbers

and

If  $\mathbb{E}X < \infty$ ,  $\{X_i\}$  i.i.d.

then  $\bar{X}_n \xrightarrow{P.} \mathbb{E}X$

Continuous Mapping Theorem.

if a sequence of random variables  
converge in probability:  $Y_n \xrightarrow{P.} c$ .

$f(\cdot)$  is a continuous function, then

$f(Y_n) \xrightarrow{P.} f(c)$

## ST102 Exercise 15

In this exercise you will practise aspects of asymptotic distributions and interval estimation. Question 1 requires you to obtain the asymptotic sampling distribution of the maximum likelihood estimator of the Bernoulli parameter,  $\pi$ . Question 2 uses the sample standard deviation for determining a confidence interval for the population mean. In Question 3, think about using a two-point (i.e. Bernoulli) distribution. Finally, Question 4 requires you to evaluate whether ' $\sigma$ ' or the sample standard deviation is appropriate to calculate a confidence interval.

Your answers to this problem set should be submitted as a pdf file upload to Moodle, *as directed by your class teacher*. It will be covered by your class teacher in your fifteenth class, which will take place in the week commencing Monday 19 February 2024.

- 1.\* Let  $\{X_1, X_2, \dots, X_n\}$  be a random sample from a Bernoulli distribution with parameter  $\pi$ , where  $0 < \pi < 1$ . By computing Fisher information, determine the asymptotic distribution of the maximum likelihood estimator of  $\pi$ , which is  $\hat{\pi} = \bar{X}$ .
2. In a survey of students, the number of hours per week of private study is recorded. For a random sample of 22 students, the sample mean is 19.1 hours and the sample standard deviation is 3.8 hours. Treat the data as a random sample from a normal distribution.
  - (a) Find a 90% confidence interval for the mean number of hours per week of private study in the student population.
  - (b) Recompute your confidence interval in the case that the sample size is, in fact, 121, but the sample mean and sample standard deviation values are unchanged. Comment on the two intervals.
- 3.\* In a study of consumers' views on guarantees for new products, 410 out of a random sample of 475 consumers agreed with the statement: '*Product guarantees are worded more for lawyers to understand than to be easily understood by consumers.*'
  - (a) Find an *approximate* 95% confidence interval for the population proportion of consumers agreeing with this statement.
  - (b) Would a 99% confidence interval for the population proportion be wider or narrower than that found in (a)? Explain your answer.
4. Suppose a random survey of 400 first-time home buyers finds that the sample mean of annual household income is £36,000 and the sample standard deviation is £17,000.
  - (a) An economist believes that the 'true' standard deviation is  $\sigma = £12,000$ . Based on this assumption, find an *approximate* 90% confidence interval for  $\mu$ , i.e. for the average annual household income of all first-time home buyers.

- 1.\* Let  $\{X_1, X_2, \dots, X_n\}$  be a random sample from a Bernoulli distribution with parameter  $\pi$ , where  $0 < \pi < 1$ . By computing Fisher information, determine the asymptotic distribution of the maximum likelihood estimator of  $\pi$ , which is  $\hat{\pi} = \bar{X}$ .

Thm. maximum likelihood estimators are asymptotically normal.

$$X \sim \text{Ber}(\pi)$$

$$f(x; \pi) = \pi^x (1-\pi)^{1-x} \quad \text{for } x=0, 1$$

$$\log f(x; \pi) = x \log \pi + (1-x) \log (1-\pi)$$

$$\frac{d}{d\pi} \log f(x; \pi) = \frac{x}{\pi} - \frac{1-x}{1-\pi}$$

$$\frac{d^2}{d\pi^2} \log f(x; \pi) = -\frac{x}{\pi^2} - \frac{1-x}{(1-\pi)^2}$$

$$I(\pi) = \mathbb{E}[-\frac{d^2}{d\pi^2} \log f(x; \pi)] = \frac{1}{\pi^2} \mathbb{E}X + \frac{1}{(1-\pi)^2} (1-\mathbb{E}X) = \frac{1}{\pi} + \frac{1}{1-\pi} = \frac{1}{\pi(1-\pi)}$$

Apply the asymptotic normality of mle, we have:  $(n \cdot I(\pi))^{-1}$

$$\hat{\pi}_{MLE} - \pi \xrightarrow{d.} N(0, \underbrace{(n \cdot I(\pi))^{-1}}_{\text{as } n \rightarrow \infty})$$

$$\text{i.e. } \hat{\pi}_{MLE} - \pi \xrightarrow{d.} N(0, \frac{\pi(1-\pi)}{n}) \quad \text{as } n \rightarrow \infty$$

2. In a survey of students, the number of hours per week of private study is recorded. For a random sample of 22 students, the sample mean is 19.1 hours and the sample standard deviation is 3.8 hours. Treat the data as a random sample from a normal distribution.

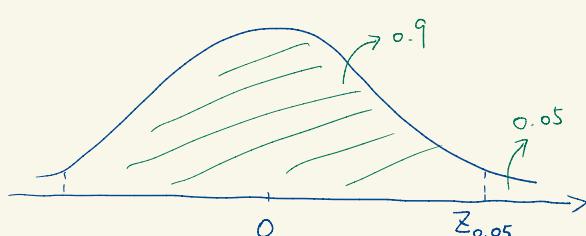
- (a) Find a 90% confidence interval for the mean number of hours per week of private study in the student population.
- (b) Recompute your confidence interval in the case that the sample size is, in fact, 121, but the sample mean and sample standard deviation values are unchanged. Comment on the two intervals.

$$X \sim N(\mu, \sigma^2)$$

"finding confidence interval"

(a) i.e. find  $\hat{\theta}_1, \hat{\theta}_2$  such that  $P(\hat{\theta}_1 \leq \bar{X} \leq \hat{\theta}_2) = 0.9$

$$\bar{X}_n - \mathbb{E}\bar{X} \sim N(0, \frac{\sigma^2}{n}) \Rightarrow \frac{\bar{X} - \mathbb{E}\bar{X}}{\sigma/\sqrt{n}} \sim N(0, 1)$$



To find Fisher information.

$$\frac{\partial}{\partial \theta} \log f(x; \theta)$$

$$\frac{\partial^2}{\partial \theta^2} \log f(x; \theta)$$

$$\mathbb{E}\left[ -\frac{\partial^2}{\partial \theta^2} \log f(x; \theta) \right] =: I(\theta)$$

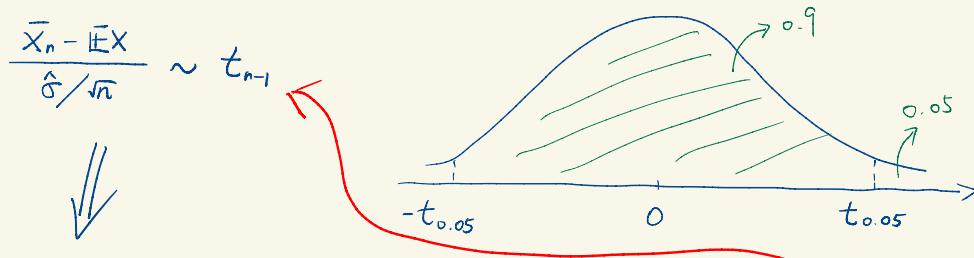
To calculate MLE.

$$l(\theta; x_1, \dots, x_n) = \sum_{i=1}^n \log f(x_i; \theta)$$

$$\frac{\partial l}{\partial \theta} = 0 \Rightarrow \hat{\theta}^{MLE}$$

However,  $\sigma$  is unknown. we estimate it first.

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



$$P(-t_{0.05} \leq \frac{\bar{X}_n - \bar{E}X}{\hat{\sigma}/\sqrt{n}} \leq t_{0.05}) = 0.9$$

The event is:  $\left| \frac{\bar{X}_n - \bar{E}X}{\hat{\sigma}/\sqrt{n}} \right| \leq t_{0.05}$

$$\text{i.e. } \bar{X}_n - \frac{\hat{\sigma}}{\sqrt{n}} \cdot t_{0.05} \leq \bar{E}X \leq \bar{X}_n + \frac{\hat{\sigma}}{\sqrt{n}} \cdot t_{0.05}$$

$n=22$ , thus  $t_{0.05}$  refers to  $t_{21, 0.05} = 1.721$

① use three cases.  
to draw conclusion  
on distribution.

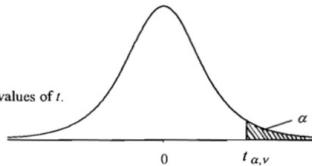
②

length of the  
interval.  $\frac{2\hat{\sigma}}{\sqrt{n}} \cdot t_{0.05}$

The values of  $t$  are obtained by solution of the equation:

$$\alpha = \Gamma[\frac{1}{2}(\nu+1)] / [\Gamma(\frac{1}{2}\nu)]^{-1} (\nu\pi)^{-1/2} \int_t^\infty (1+x^2/\nu)^{-(\nu+1)/2} dx$$

Note: The tabulation is for one tail only, that is, for positive values of  $t$ .  
For  $|t|$  the column headings for  $\alpha$  should be doubled.

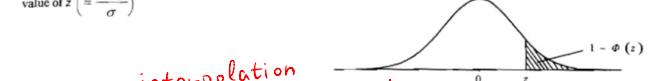


$\alpha =$	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
v = 1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
$\infty$	1.282	1.645	1.960	2.326	2.576	3.090	3.291

$t_n \xrightarrow{d} N(0, 1)$

Table 3 Areas in Upper Tail of the Normal Distribution

The function tabulated is  $1 - \Phi(z)$  where  $\Phi(z)$  is the cumulative distribution function of a standardised Normal variable,  $z$ .  
Thus  $1 - \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-x^2/2} dx$  is the probability that a standardised Normal variate selected at random will be greater than a value of  $z$  ( $= \frac{x - \mu}{\sigma}$ )



$\frac{x-\mu}{\sigma}$	.00	.01	.02	.03	<u>.04</u>	<u>.05</u>	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2809	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2098	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0750	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	<u>.0505</u>	<u>.0495</u>	.0485	.0475	.0465	.0455
1.7	.0446	.0434	.0427	.0418	.0409	<u>.0401</u>	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0244	.0239	.0233	
2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
2.1	.01786	.01745	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00154	.00149	.00144	.00139	
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00018	.00017	.00016	
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.000108	.000104	.000100	.000096	.000092	.000088	.000085	.000082	.000078	.000075
3.8	.000072	.000069	.000067	.000064	.000062	.000059	.000057	.000054	.000052	.000050
3.9	.000048	.000046	.000044	.000042	.000041	.000039	.000037	.000036	.000034	.000033
4.0	.000032									

5.0 → 0.000 000 286 7      5.5 → 0.000 000 019 0      6.0 → 0.000 000 001 0

This table is taken from Table III of Fisher & Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, reprinted by permission of Addison Wesley Longman Ltd. Also from Table 12 of *Biometrika Tables for Statisticians*, Volume 1, by permission of Oxford University Press and the Biometrika Trustees.

(b) change  $n$  from 22 to 121.

One would expect the range of the confidence interval is smaller.  
 The intuition behind is that, with a larger sample size, we are more confident to conclude that the population mean is in a smaller region. (the estimate is more accurate)

$$t_{120, 0.05} = 1.658, t_{21, 0.05} = 1.721$$

Compare with normal.  $Z_{0.05} = 1.645 \approx t_{120, 0.05}$

No Normal assumption. Agree or not is Bernoulli

3.\* In a study of consumers' views on guarantees for new products, 410 out of a random sample of 475 consumers agreed with the statement: 'Product guarantees are worded more for lawyers to understand than to be easily understood by consumers.'

475 is large  
 ↓  
 apply ③.

- (a) Find an approximate 95% confidence interval for the population proportion of consumers agreeing with this statement.
- (b) Would a 99% confidence interval for the population proportion be wider or narrower than that found in (a)? Explain your answer.

$$(a) X_i = \begin{cases} 1 & \text{if } i\text{-th consumer agree with ...} \\ 0 & \text{--- --- disagree ---} \end{cases}$$

then  $X_i \sim \text{Bernoulli}(\pi)$

$$\text{apply ③: } \frac{\bar{X}_n - \mathbb{E}X}{\hat{\sigma}/\sqrt{n}} \xrightarrow{\text{app.}} N(0, 1)$$

$\hat{\pi}^2 = \pi(1-\pi)$ . By LLN,  $\hat{\pi} = \bar{X}_n$  is consistent,  
 $\Downarrow$  continuous mapping theorem.

$\hat{\pi}^2 = \hat{\pi}(1-\hat{\pi})$  is consistent

$$P(-Z_{0.025} \leq \frac{\bar{X}_n - \mathbb{E}X}{\hat{\sigma}/\sqrt{n}} \leq Z_{0.025}) \approx 0.95$$

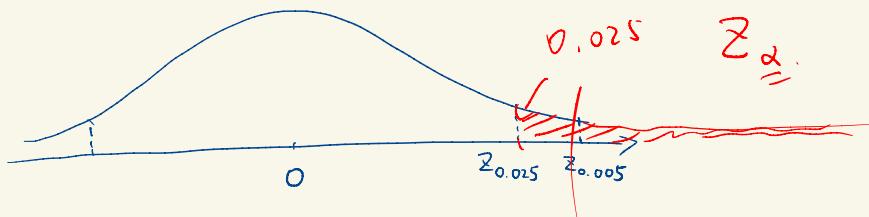
$$\left( \bar{X}_n - Z_{0.025} \cdot \frac{\hat{\sigma}}{\sqrt{n}}, \bar{X}_n + Z_{0.025} \cdot \frac{\hat{\sigma}}{\sqrt{n}} \right)$$

$$\text{numerically, } \bar{X}_n = \frac{410}{475}, \quad \hat{\sigma} = \frac{410 \times 65}{475^2}$$

$$Z_{0.025} = 1.96, n = 475$$

knowing without proof is helpful.  
 for any continuous function  
 $f$ .  $f(\hat{\pi})$  is consistent for  $f(\pi)$   
 e.g.  $\hat{\pi}^2 \xrightarrow{P} \pi^2$   
 $e^{\hat{\pi}} \xrightarrow{P} e^\pi$

(b) a 99% CI is wider: one has to include more potential values for  $\bar{X}_n - \pi$  to take.



$$(\bar{X}_n - z_{0.005} \cdot \frac{\hat{\sigma}}{\sqrt{n}}, \bar{X}_n + z_{0.005} \cdot \frac{\hat{\sigma}}{\sqrt{n}})$$

compared to

$$(\bar{X}_n - z_{0.025} \cdot \frac{\hat{\sigma}}{\sqrt{n}}, \bar{X}_n + z_{0.025} \cdot \frac{\hat{\sigma}}{\sqrt{n}})$$

4. Suppose a random survey of 400 first-time home buyers finds that the sample mean of annual household income is £36,000 and the sample standard deviation is £17,000.
- An economist believes that the 'true' standard deviation is  $\sigma = \text{£}12,000$ . Based on this assumption, find an *approximate* 90% confidence interval for  $\mu$ , i.e. for the average annual household income of all first-time home buyers.
  - Without the assumption that  $\sigma$  is known, find an *approximate* 90% confidence interval for  $\mu$ .
  - Are the two confidence intervals very different? Which one would you trust more, and why?

(a)  $\sigma$  is known, use ①. 
$$\frac{\bar{X}_n - \mathbb{E}X}{\sigma/\sqrt{n}} \stackrel{\text{app.}}{\sim} N(0,1)$$

$$(\bar{X}_n - Z_{0.05} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X}_n + Z_{0.05} \cdot \frac{\sigma}{\sqrt{n}})$$

(b)  $\sigma$  is unknown,  $n=400$  is large. No normality, use ③.

$$\frac{\bar{X}_n - \mathbb{E}X}{\hat{\sigma}/\sqrt{n}} \stackrel{\text{app.}}{\sim} N(0,1)$$

$$(\bar{X}_n - Z_{0.05} \cdot \frac{\hat{\sigma}}{\sqrt{n}}, \bar{X}_n + Z_{0.05} \cdot \frac{\hat{\sigma}}{\sqrt{n}})$$

(c) I trust the one in (b) more, since the 'sample sd' comes from a sample, whilst the economists' belief comes from nowhere.

If  $\sigma = \text{£}12000$  can be elaborated, and sounds more reasonable than  $\hat{\sigma} = \text{£}17000$ , I might change my mind.