STA260: PROBABILITY AND STATISTICS II
SPRING 2021

TUTORIAL 1 (TUT9101)

MAY 5, 2021

YUEHUAN HE

UNIVERSITY OF TORONTO

Problem 1

7.9 Refer to Example 7.2. The amount of fill dispensed by a bottling machine is normally distributed with $\sigma = 1$ ounce. If $n = 9$ bottles are randomly selected from the output of the machine, we found that the probability that the sample mean will be within .3 ounce of the true mean is .6318. Suppose that $\overline{Y}$ is to be computed using a sample of size n.

a If $n = 16$, what is $P(|\overline{Y} - \mu| \leq .3)$?

b Find $P(|\overline{Y} - \mu| \leq .3)$ when $\overline{Y}$ is to be computed using samples of sizes $n = 25$, $n = 36$, $n = 49$, and $n = 64$.

c What pattern do you observe among the values for $P(|\overline{Y} - \mu| \leq .3)$ that you observed for the various values of n?

d Do the results that you obtained in part (b) seem to be consistent with the result obtained in Example 7.3?

$\boxed{\frac{1}{\sqrt{n}}}$ standard error

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

$\overline{Y} \sim N(\mu, \frac{\sigma^2}{n})$

$\overline{Y} - \mu \sim N(0, \frac{\sigma^2}{n})$

$Z \sim N(0,1)$

$Y_1 \cdots Y_n$ are the output. $\{Y_i\}_{i=1}^n \overset{iid}{\sim} N(\mu, \sigma^2)$ $\sigma = 1$.

from ~~lN~~ notes problem 7.2.4

$\overline{Y} = \frac{1}{n}(Y_1 + Y_2 + \cdots + Y_n) = \frac{1}{n}\sum_{i=1}^{n}$

$\overline{Y} \sim N(\mu, \frac{\sigma^2}{n})$.

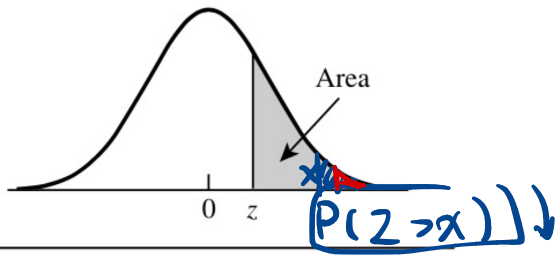
a. $P(|\overline{Y} - \mu| \leq 0.3)$

$= P\left(\frac{|\overline{Y} - \mu|}{\left(\frac{\sigma}{\sqrt{n}}\right)} \leq \frac{0.3}{\left(\frac{\sigma}{\sqrt{n}}\right)}\right)$ $= P\left(|Z| \leq \frac{0.3}{\sigma/\sqrt{n}}\right)$

Normal Distribution

$\boxed{\sigma = 1}$

Table 4 Normal Curve Areas
Standard normal probability in right-hand tail
(for negative values of z, areas are found by symmetry)



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
	Second decimal place of z									
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	.2810	.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	.2090	.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

a. $P(|\bar{Y} - \mu| \le 0.3)$

$= P\left(|Z| \le \frac{0.3}{\sigma} \sqrt{n}\right)$ 4. → 16

$= P(|Z| \le 1.2)$

$= P(-1.2 \le Z \le 1.2)$

$= 1 - 2\,\underline{P(Z > 1.2)}$

$= 1 - 2 \times 0.1151$

$= \underline{0.7698}$

b. $P_n = P(|\bar{Y}-\mu| \le 0.3) = P\left(|Z| \le \frac{0.3}{\sigma}\sqrt{n}\right)$.

$= 1 - 2 \times P\left(Z > \boxed{\frac{0.3}{\sigma}\sqrt{n}}\right)$

$n=25$ $\quad P_{25} = 1 - 2P(Z > 1.5) = 0.866$

$n=36$ $\quad P_{36} = 1 - 2P(Z > 1.8) = 0.928$

$n=49$ $\quad P_{49} = 1 - 2P(Z > 2.1) = 0.964$

$n=64$ $\quad P_{64} = 1 - 2P(Z > 2.4) = 0.984$.

c. $n \uparrow \quad P_n \uparrow$.

$P\left(Z > \boxed{\frac{0.3}{\sigma}}\sqrt{n}\right)$. $\uparrow \quad \downarrow$

$(1 - 2 \times P_{\downarrow}) \uparrow$

EXAMPLE 7.3 Refer to Example 7.2. How many observations should be included in the sample if we wish $\overline{Y}$ to be within .3 ounce of μ with probability .95?

d Do the results that you obtained in part (b) seem to be consistent with the result obtained in Example 7.3?

eg 7.3. n? to het $P(|\bar{Y}-\mu| \le 0.3) = 0.95$.

$P(-1.96 \le Z \le 1.96) = 0.95$

$$P\left(-\frac{0.3\sqrt{n}}{\sigma} \le Z \le \frac{0.3\sqrt{n}}{\sigma}\right) = 0.95.$$

$$\frac{0.3\sqrt{n}}{\sigma} \geq 1.96 \qquad \sqrt{n} \geq \frac{1.96}{0.03} \qquad n \geq 43$$

7.13 The Environmental Protection Agency is concerned with the problem of setting criteria for the amounts of certain toxic chemicals to be allowed in freshwater lakes and rivers. A common measure of toxicity for any pollutant is the concentration of the pollutant that will kill half of the test species in a given amount of time (usually 96 hours for fish species). This measure is called LC50 (lethal concentration killing 50% of the test species). In many studies, the values contained in the natural logarithm of LC50 measurements are normally distributed, and, hence, the analysis is based on ln(LC50) data.

Studies of the effects of copper on a certain species of fish (say, species A) show the variance of ln(LC50) measurements to be around .4 with concentration measurements in milligrams per liter. If $n = 10$ studies on LC50 for copper are to be completed, find the probability that the sample mean of ln(LC50) will differ from the true population mean by no more than .5.

μ

$Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$. $\sigma^2 = 0.4$. $\quad \bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. $0.2^2 = 0.04$

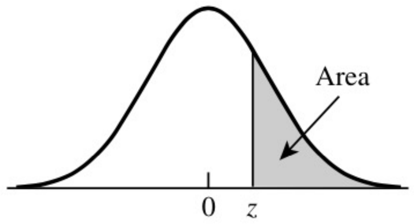
$$P(|\bar{Y} - \mu| \leq 0.5) = P\left(\boxed{\frac{|\bar{Y} - \mu|}{\frac{\sigma}{\sqrt{n}}}} \leq \frac{0.5}{\sqrt{\frac{\sigma^2}{n}}} \right) \qquad \sqrt{\frac{0.4}{10}} = \sqrt{0.04} = 0.2$$

$$= P(\boxed{|Z|} \leq 2.5)$$

$$= 1 - 2P(Z > 2.5) = 0.9876.$$

Normal Distribution

Table 4 Normal Curve Areas
Standard normal probability in right-hand tail
(for negative values of z, areas are found by symmetry)



	Second decimal place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.5	.0062									
2.6	.0047									
2.7	.0035									
2.8	.0026									
2.9	.0019									

7.21 Refer to Exercise 7.13. Suppose that $n = 20$ observations are to be taken on ln(LC50) measurements and that $\sigma^2 = 1.4$. Let S^2 denote the sample variance of the 20 measurements.

a Find a number b such that $P(S^2 \le b) = .975$.

b Find a number a such that $P(a \le S^2) = .975$.

c If a and b are as in parts (a) and (b), what is $P(a \le S^2 \le b)$?

theorem 7.2.17 $Y_1 \cdots Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$. $\boxed{\frac{(n-1)S^2}{\sigma^2}} \sim \boxed{\chi^2_{(n-1)}}$.

$$a.\ P(S^2 \le b) = P\left(\frac{(n-1)S^2}{\sigma^2} \le \frac{(n-1)\cdot b}{\sigma^2}\right) = 0.975$$

~~follows~~ follows $\chi^2_{(n-1)}$.

7.21 Refer to Exercise 7.13. Suppose that $n = 20$ observations are to be taken on ln(LC50) measurements and that $\sigma^2 = 1.4$. Let S^2 denote the sample variance of the 20 measurements.

a Find a number b such that $P(S^2 \le b) = .975.$

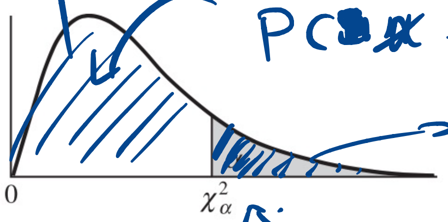
0.975

$X \sim \chi^2(n-1)$

19

$$\boxed{\frac{(n-1)\,b}{\sigma^2}} = 32.8523$$

chi square-Distribution



$P(X \le a) = $ ___

$1 - 0.975$

0.025

$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$	df
22.3072	24.9958	27.4884	30.5779	32.8013	15
23.5418	26.2962	28.8454	31.9999	34.2672	16
24.7690	27.5871	30.1910	33.4087	35.7185	17
25.9894	28.8693	31.5264	34.8053	37.1564	18
27.2036	30.1435	$\boxed{32.8523}$	36.1908	38.5822	19

$$\frac{19\,b}{1.4} = 32.8523$$

$$b = 2.42$$

b Find a number a such that $P(a \le S^2) = .975$.

$$P(a \le S^2) = P\left(\boxed{\frac{(n-1)a}{\sigma^2}} \le \frac{(n-1)S^2}{\sigma^2}\right)$$

$\chi^2_{(n-1)}$.

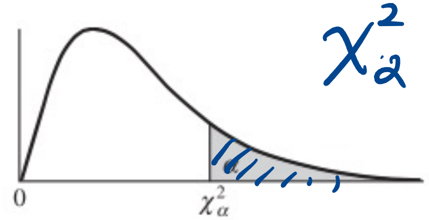
19

$$\frac{(n-1)a}{\sigma^2} = 8.90655$$

1.4

$$a = 0.656.$$

chi square-Distribution



χ^2_α

df	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509

c If a and b are as in parts (a) and (b), what is $P(a \le S^2 \le b)$?

$$P(a \le S^2 \le b) = 1 - P(S^2 > b) - P(S^2 < a)$$

0.025

0.975

$$= 1 - (1 - P(S^2 \le b))$$

0.025

$$- (1 - P(a \le S^2))$$

$$= 0.95$$

chi square-Distribution