STA260: PROBABILITY AND STATISTICS II
SPRING 2021

TUTORIAL 3 (TUT9101)

MAY 12, 2021

YUEHUAN HE

UNIVERSITY OF TORONTO

7.49 The length of time required for the periodic maintenance of an automobile or another machine usually has a mound-shaped probability distribution. Because some occasional long service times will occur, the distribution tends to be skewed to the right. Suppose that the length of time required to run a 5000-mile check and to service an automobile has mean 1.4 hours and standard deviation .7 hour. Suppose also that the service department plans to service 50 automobiles per 8-hour day and that, in order to do so, it can spend a maximum average service time of only 1.6 hours per automobile. On what proportion of all workdays will the service department have to work overtime? $(\bar{Y} > 1.6)$

$$\frac{8}{1.6} = 5$$

$\mu = 1.4$, $\sigma = 0.7$, $n = 50$

$$P(\bar{Y} > 1.6) = P\left(\frac{\bar{Y} - 1.4}{\frac{0.7}{\sqrt{50}}} > \frac{1.6 - 1.4}{\frac{0.7}{\sqrt{50}}}\right)$$

$$= P(Z > 2.02)$$

$$= 0.0217$$

UNIVERSITY OF TORONTO

7.50 Shear strength measurements for spot welds have been found to have standard deviation 10 pounds per square inch (psi). If 100 test welds are to be measured, what is the approximate probability that the sample mean will be within 1 psi of the true population mean?

7.51 Refer to Exercise 7.50. If the standard deviation of shear strength measurements for spot welds is 10 psi, how many test welds should be sampled if we want the sample mean to be within 1 psi of the true mean with probability approximately .99?

$\sigma = 10$, $n = ?$

$P(|\bar{Y} - \mu| \le 1) = 0.99$

$\Downarrow$

$P\left(\dfrac{|\bar{Y} - \mu|}{\frac{\sigma}{\sqrt{n}}} \le \dfrac{1}{\frac{\sigma}{\sqrt{n}}}\right) = 0.99$

$\Downarrow$

$P\left(|Z| \le \boxed{\dfrac{\sqrt{n}}{10}}\right) = 0.99$

$$\begin{cases} P(Z < -2.57) = 0.005 \\ P(Z > 2.57) = 0.005 \end{cases}$$

2.57

$\therefore \dfrac{\sqrt{n}}{10} = \boxed{Z_{0.005}}$

$\Downarrow$

$n \simeq 664$

7.57 Twenty-five heat lamps are connected in a greenhouse so that when one lamp fails, another takes over immediately. (Only one lamp is turned on at any time.) The lamps operate independently, and each has a mean life of 50 hours and standard deviation of 4 hours. If the greenhouse is not checked for 1300 hours after the lamp system is turned on, what is the probability that a lamp will be burning at the end of the 1300-hour period?

$n = 25 \quad Y_1 + Y_2 \cdots Y_{25}$

$$\frac{Y_1 + Y_2 + \cdots + Y_{25}}{25} \qquad > 1300$$

$\bar{Y}$

$$\bar{Y} = \frac{1}{n}\left[Y_1 + \cdots + Y_n\right]$$

7.57 Twenty-five heat lamps are connected in a greenhouse so that when one lamp fails, another takes over immediately. (Only one lamp is turned on at any time.) The lamps operate independently, and each has a mean life of 50 hours and standard deviation of 4 hours. If the greenhouse is not checked for 1300 hours after the lamp system is turned on, what is the probability that a lamp will be burning at the end of the 1300-hour period?

$\mu = 50, \ \sigma = 4, \ n = 25$

$$P(Y_1 + \cdots + Y_{50} \geq 1300) = P\left(\bar{Y} \geq \frac{1300}{n}\right)$$

$$= P\left(\frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} \geq \boxed{\frac{\frac{1300}{n} - \mu}{\frac{\sigma}{\sqrt{n}}}}\right)$$

$$= P(Z \geq 2.5)$$

$$= 0.0062$$

UNIVERSITY OF TORONTO

Problem 4

7.58 Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ are independent random samples from populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. Show that the random variable

$$U_n = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

satisfies the conditions of Theorem 7.4 and thus that the distribution function of U_n converges to a standard normal distribution function as $n \to \infty$. [*Hint:* Consider $W_i = X_i - Y_i$, for $i = 1, 2, \ldots, n$.]

The Central Limit Theorem

Theorem 7.2.32 (Theorem 7.4 (page 372) of [3]:) Let $Y_1, Y_2, \ldots, Y_n$ be independent and indentically distributed random variables with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2 < \infty$. Define

$$U_n = \frac{\sum_{i=1}^n Y_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$$

That is,

$$\lim_{n\to\infty} P(U_n \le u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \qquad \text{for all } u$$

7.58 Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ are independent random samples from populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. Show that the random variable

$$U_n = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

satisfies the conditions of Theorem 7.4 and thus that the distribution function of U_n converges to a standard normal distribution function as $n \to \infty$. [*Hint:* Consider $W_i = X_i - Y_i$, for $i = 1, 2, \ldots, n$.]

The Central Limit Theorem

Theorem 7.2.32 (Theorem 7.4 (page 372) of [3]:) Let $Y_1, Y_2, \ldots, Y_n$ be independent and indentically distributed random variables with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2 < \infty$. Define

$$U_n = \boxed{\frac{\sum_{i=1}^n Y_i - n\mu}{\sigma\sqrt{n}}} = \boxed{\frac{\bar{Y} - \mu}{\sigma/\sqrt{n}}} \xrightarrow{d} Z \sim N(0,1)$$

That is,

$$\lim_{n\to\infty} P(U_n \le u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \qquad \text{for all } u$$

replace Y with W

$W_i = X_i - Y_i \Rightarrow \overline{W} = \overline{X} - \overline{Y}$

$E(W_i) = \mu_1 - \mu_2$

$Var(W_i) = Var(X_i) + Var(Y_i) = \sqrt{\sigma_1^2 + \sigma_2^2}$

$\Downarrow$

$\sigma_W = \sqrt{\sigma_1^2 + \sigma_2^2}$

$\overline{W} - E(W_i)$

$$U_n = \frac{\overline{W} - (\mu_1 - \mu_2)}{\sqrt{\frac{(\sigma_1^2 + \sigma_2^2)}{n}}} \xrightarrow{d} N(0,1)$$

7.73 An airline finds that 5% of the persons who make reservations on a certain flight do not show up for the flight. If the airline sells 160 tickets for a flight with only 155 seats, what is the probability that a seat will be available for every person holding a reservation and planning to fly?

$n = 160, \quad p = 0.95$

(normal approximation)

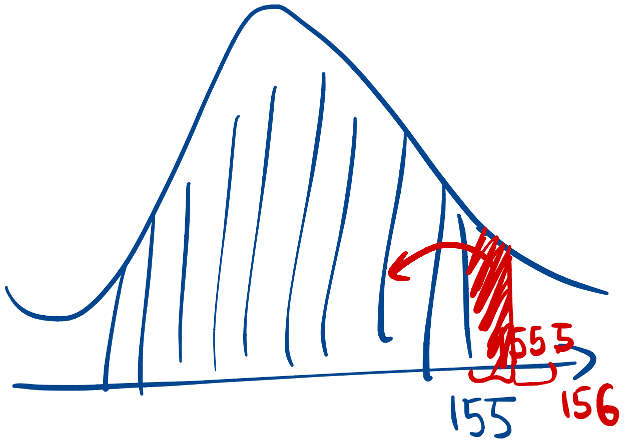
$$Y \sim Bin(n, p)$$

$$Y \approx N(np, npq)$$

$$P(Y \le 155) \approx P\left(\frac{Y - np}{\sqrt{npq}} \le \frac{155.5 - np}{\sqrt{npq}}\right)$$

$$= P(Z \le 1.27)$$

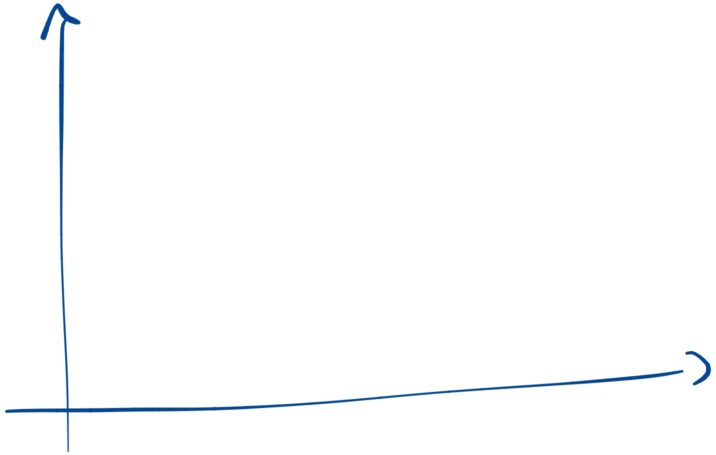
$$= 0.8980$$



(continuity correction)

$$P(Y \le b) \approx P(Y \le b + 0.5).$$

$$P(Y \le b) \approx P(Y \le b + 0.5)$$



Problem 6

7.77 The manager of a supermarket wants to obtain information about the proportion of customers who dislike a new policy on cashing checks. How many customers should he sample if he wants the sample fraction to be within .15 of the true fraction, with probability .98?

$n = ?$

$$P\left(\left|\frac{Y}{n} - p\right| \le 0.15\right) = 0.98$$

$Y \approx N(np, npq)$

$\frac{1}{n}Y \approx N(p, \frac{1}{n}pq)$

$$\Downarrow$$

$$P\left(\frac{\left|\frac{Y}{n} - p\right|}{\sqrt{\frac{1}{n}pq}} \le \frac{0.15}{\sqrt{\frac{1}{n}pq}}\right) = 0.98$$

$\max pq = p(1-p) = \boxed{\frac{1}{4}}$

$n \le 61$

$$P\left(|Z| \le \frac{0.15}{\sqrt{\frac{1}{n}pq}}\right) = 0.98$$

$$\frac{0.15}{\sqrt{\frac{1}{n}pq}} = Z_{0.01} = 2.33 \Rightarrow \sqrt{n} = \frac{2.33}{0.15}\boxed{\sqrt{pq}} \le 7.76$$

$\sqrt{\frac{1}{4}}$

UNIVERSITY OF TORONTO

Problem 7

7.81 A lot acceptance sampling plan for large lots specifies that 50 items be randomly selected and that the lot be accepted if no more than 5 of the items selected do not conform to specifications.

a What is the approximate probability that a lot will be accepted if the true proportion of nonconforming items in the lot is .10?

b Answer the question in part (a) if the true proportion of nonconforming items in the lot is .20 and .30.

$P(Y \le b)$

$\downarrow$

$b + 0.5$

continuity correction

a. $p = 0.1$

$n = 50$

$$P(\text{accepted}) = P(Y \le 5) \approx P(Y \le 5.5)$$

$$= P\left(\frac{Y - np}{\sqrt{npq}} \le \frac{5.5 - np}{\sqrt{npq}}\right)$$

$$= P\left(Z \le \frac{0.5}{\sqrt{50 \times 0.1 \times 0.9}}\right)$$

$$= P(Z \le 0.236)$$

$$= 0.5948$$

7.81 A lot acceptance sampling plan for large lots specifies that 50 items be randomly selected and that the lot be accepted if no more than 5 of the items selected do not conform to specifications.

a What is the approximate probability that a lot will be accepted if the true proportion of nonconforming items in the lot is .10?

b Answer the question in part (a) if the true proportion of nonconforming items in the lot is .20 and .30.

b. $p = 0.2$

$p = 0.3$

$$P(\text{accepted}) \approx P\left(\frac{Y - np}{\sqrt{npq}} \le \frac{5.5 - np}{\sqrt{npq}}\right)$$

$$\approx \begin{cases} 0.056 & p = 0.2 \\ 0.002 & p = 0.3 \end{cases}$$

Problem 8

7.84 Just as the difference between two sample means is normally distributed for large samples, so is the difference between two sample proportions. That is, if Y_1 and Y_2 are independent binomial random variables with parameters (n_1, p_1) and (n_2, p_2), respectively, then $(Y_1/n_1) - (Y_2/n_2)$ is approximately normally distributed for large values of n_1 and n_2.

a Find $E\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right)$.

b Find $V\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right)$.

$Y_1 \sim Bin(n_1, p_1)$

$Y_2 \sim Bin(n_2, p_2)$.

7.84 Just as the difference between two sample means is normally distributed for large samples, so is the difference between two sample proportions. That is, if Y_1 and Y_2 are independent binomial random variables with parameters (n_1, p_1) and (n_2, p_2), respectively, then $(Y_1/n_1) - (Y_2/n_2)$ is approximately normally distributed for large values of n_1 and n_2.

a Find $E\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right)$.

b Find $V\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right)$.

$$a \quad E\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right) = \frac{E(Y_1)}{n_1} - \frac{E(Y_2)}{n_2} = \frac{n_1 P_1}{n_1} - \frac{n_2 P_2}{n_2} = P_1 - P_2$$

$$b. \quad V\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right) = \frac{V(Y_1)}{n_1^2} + \frac{V(Y_2)}{n_2^2} = \frac{n_1 P_1 q_1}{n_1^2} + \frac{n_2 P_2 q_2}{n_2^2}$$

$$= \frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}$$

$n = 100$

7.87 The times to process orders at the service counter of a pharmacy are exponentially distributed with mean 10 minutes. If 100 customers visit the counter in a 2-day period, what is the probability that at least half of them need to wait more than 10 minutes?

$(Y \geq 50)$

$Y \sim Bin(n, p)$

$$\boxed{p} = \int_{10}^{\infty} \frac{1}{10} e^{-\frac{y}{10}} dy$$

$$= 0.3679$$

$n = 100, \quad p \simeq 0.3679$

continuity correction

49.5

$$P(Y \geq 50) \approx P(Y \geq 50 - 0.5)$$

$$= P\left(Z \geq \boxed{\frac{49.5 - np}{\sqrt{npq}}}\right)$$

$$= P(Z \geq 2.636)$$

$$= 0.0042$$

$P(Y \leq b)$.

$P(a \leq Y)$.

$\parallel$

$P(Y \geq a)$.

THANK YOU FOR ATTENDING TODAY’S TUTORIAL

Feedbacks: yuehuan.he@utoronto.ca