Quiz 4 review key

Stat 301

Summer 2019

(1) A service station had both self-service and full-service islands. On each island, there is a single regular unleaded pump with two hoses. Let X denote the number of hoses being used on the self-service island at a particular time, let Y denote the number of hoses on the full-service island in use at that time. The joint pmf of X and Y are in the accompanying table.

(a) Calculate the Marginal distributions of X and Y Marginal Distributions are just the row and column totals (see filled out table below).

$p_X(0) = 0.15 + 0.02 + 0.01 = 0.18, \ldots,; p_Y(0) = 0.15 + 0.06 + 0.03 = 0.24, \ldots,$

$$p_X(x) = \begin{cases} 0.18 & x = 0 \\ 0.31 & x = 1 \\ 0.51 & x = 2 \end{cases}$$

$$p_Y(y) = \begin{cases} 0.24 & y = 0 \\ 0.38 & y = 1, 2 \end{cases}$$

(b) Calculate EX, EY, $E(XY)$, VX, VY, SDX, SDY, $Cov(X,Y)$, $Corr(X,Y)$. Once $Corr(X,Y)$ is calculated, explain what it means.

$$EX = \sum xp(x) = 0(0.18) + 1(0.31) + 2(0.51) = 1.33$$

$$EY = \sum yp(y) = 0(0.24) + 1(0.38) + 2(0.38) = 1.14$$

$$VX = E(X^2) - (EX)^2 \ with \ E(X^2) = \sum x^2 p(x)$$

$$E(X^2) = 0^2(0.18) + 1^2(0.31) + 2^2(0.51) = 2.35$$

$$E(Y^2) = \sum y^2 p(y) = 0^2(0.24) + 1^2(0.38) + 2^2(0.38) = 1.9$$

$$VX = 2.35 - (1.33)^2 = 0.5811$$
$$VY = 1.9 - (1.14)^2 = 0.6004$$

$$SDX = \sqrt{VX} = \sqrt{0.5811} = 0.7622992 \quad SDY = \sqrt{0.6004} = 0.7748548$$

$$E(XY) = \sum_X \sum_Y xy \ p(x,y) = 0(0)(0.15) + 0(1)(0.02) + 0(2)(0.01) + 1(0)(0.06)$$

$$+1(1)(0.18)+1(2)(0.07)+2(0)(0.03)+2(1)(0.18)+2(2)(0.3)=1.88$$

$$Cov(X,Y)=E(XY)-(EX)(EY)=1.88-(1.33)(1.14)=0.3638$$

$$Corr(X,Y)=\frac{Cov(X,Y)}{SDX(SDY)}=\frac{0.3638}{(0.7623)(0.7774)}=0.6138924$$

With a correlation of 0.6139, that implies there is a moderate, positive linear relationship between X and Y
(c) Calculate the probability there is exactly one hose in use at each station

$$P(X=1,Y=1)=0.18$$

(d) Calculate the probability there is at most one hose in use at each station

$$P(X\leq 1,Y\leq 1)=0.15+0.02+0.06+0.18=0.41$$

(e) Given that two hoses are in use at the self-service island, what is the probability there is at most one hose in use on the full-service island?

$$P(Y\leq 1|X=2)=\frac{P(Y\leq 1,X=2)}{P(X=2)}=\frac{0.03+0.18}{0.51}=0.4117647$$

(f) Given that two hoses are in use at the full-service island, what is the probability there is at most one hose in use on the self-service island?

$$P(X\leq 1|Y=2)=\frac{P(X\leq 1,Y=2)}{P(Y=2)}=\frac{0.01+0.07}{0.38}=0.2105263$$

		Y			
	$p(x,y)$	0	1	2	$p_X(x)$
	0	0.15	0.02	0.01	0.18
X	1	0.06	0.18	0.07	0.31
	2	0.03	0.18	0.30	0.51
	$p_Y(y)$	0.24	0.38	0.38	1

(2) Ithaca, NY is located in upstate New York and averages around 37" of rain each year, with standard deviation of 3.25“. Rainfall in Ithaca, NY tends to follow an approximately normal distribution. Say that a student attends Conrnell University in Ithaca and is there for 4 years.

(a) Define the Central Limit Theorem.
The sampling distribution of the sample mean will be approximately normal with mean μ and standard deviation (also known as standard error) $\frac{\sigma}{\sqrt{n}}$, provided the sample size, n, is large ($n\geq 30$)

(b) Describe the sampling distribution of the sample mean of rainfall in Ithaca, NY. Include the mean of the sampling distribution of the sample mean and the standard deviation of the sampling distribution of the sample mean (standard error).
The distribution should be approximately normal with mean 37 and standard error $=\frac{3.25}{\sqrt{4}}=1.625 \therefore \overline{X}\sim N(37,1.625)$

(c) What is the probability that during the 4 years, we see an average less than 32“?

Now use $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $P(\bar{x} < 32) = P(Z < \frac{32-37}{\frac{3.25}{\sqrt{4}}}) = P(Z < -3.08) = 0$

(d) What is the 93^{rd} percentile of average precipitation? $z_{top7\%} = z_{.93} = 1.48$ now solve for $\overline{X}$ where $\overline{X} = z(\sigma/\sqrt{n}) + \mu \Rightarrow \overline{X} = (1.48)(1.625) + 37 = 39.4$

(3) Ithaca, NY is located in upstate New York and averages around 37" of rain each year, with standard deviation of 3.25“. Rainfall in Ithaca, NY tends to follow an approximately normal distribution. Say that a student attends Conrnell University in Ithaca and is there for 7 years (undergraduate and graduate degrees).

(a) Describe the sampling distribution of the total rainfall in Ithaca, NY. Include the total of the sampling distribution and the standard deviation of the sampling distribution of the total (standard error)
$\hat{\tau} \sim N(n\mu, \sqrt{n}\sigma) \Rightarrow \hat{\tau} \sim N(259, 8.5986918)$

(b) What is the probability that during the 7 years, we see a total precipitation of less than 220“?
$P(\hat{\tau} < 220) = P(Z < \frac{220-259}{8.6}) = P(Z < -4.5348837) = 2.8127114 \times 10^{-6} \approx 0$

(c) What is the 10^{th} percentile for total precipitation?
$z_{10\%} = -1.28 \Rightarrow -1.28 = \frac{\hat{\tau}-259}{8.6} \Rightarrow \hat{\tau} = 247.98$"

(4) A survey of purchasing agents from 250 randomly selected industrial companies found that 25% of the buyers reported higher levels of new orders in January than in earlier months.

(a) Describe the sampling distribution of the proportion of buyers in the US with higher levles of new orders in January. Include the mean of the sampling distribution and the standard deviation of the sampling distribution (standard error)
$\hat{p} \sim N(p, \sqrt{pq/n}) \Rightarrow \hat{p} \sim N(0.25, 0.0273861)$

(b) What is the probability that the sample proportion is more than 26%?
$P(\hat{p} > 0.26) = P(Z > \frac{0.26-0.25}{0.0274}) = P(Z > 0.36) = 1 - P(Z < 0.36) = 1 - 0.6406 = 0.3594$

(c) What is the probability that the sample proportion is less than 20%?
$P(\hat{p} < 0.2) = P(Z < \frac{0.2-0.25}{0.0274}) = P(Z < -1.82) = 0.0344$