MS9001 – TEST 1 Solution AY 2020/21 S2

<table>
<tr><th colspan="2">Question No.</th><th>Solution</th></tr>
<tr><td rowspan="9">1</td><td>(a)</td><td>Population: All the animal bites in Kentucky.
Sample: 280 animal bites in Kentucky recorded in the dataset.</td></tr>
<tr><td>(b)</td><td>
<table>
<tr><td>SpeciesIDDesc</td><td>Qualitative</td><td>Nominal</td></tr>
<tr><td>GenderIdDesc</td><td>Qualitative</td><td>Nominal</td></tr>
<tr><td>Color</td><td>Qualitative</td><td>Nominal</td></tr>
<tr><td>WhereBittenDesc</td><td>Qualitative</td><td>Nominal</td></tr>
</table>
</td></tr>
<tr><td>(c)</td><td>Mean = $2,748
Median = $2,672</td></tr>
<tr><td>(d)</td><td>Non-normal
There are several bins that are of the similar height. OR
The histogram is not bell-shaped. OR
The histogram is not symmetrical about the mean.</td></tr>
<tr><td>(e)</td><td>CAT, DOG & RACCOON
15.7%, 77.9% & 6.4% respectively.</td></tr>
<tr><td>(f)</td><td>BODY & HEAD
80.7% and 19.3% respectively.</td></tr>
<tr><td>(g)</td><td>“The middle 50% of the fines imposed on racoon owners are between $ 1,513 and $4,153.”
IQR is $2,641.
“ The middle 50% of fines imposed on dog owners are between $1,649 and $3,943.”
IQR is $2,294.
“The middle 50% of fines imposed on cat owners are between $1,377 and $3664.”
IQR is $2,287.</td></tr>
<tr><td>(h)</td><td>
<table>
<tr><th>Variable</th><th>Total Count</th><th>Mean</th><th>Minimum</th><th>Maximum</th></tr>
<tr><td>Fine (USD)for dog bites</td><td>218</td><td>2759.0</td><td>519.0</td><td>4995.0</td></tr>
</table>
</td></tr>
<tr><td>(i)</td><td>The most common animal bites are from the dogs with the highest proportion of 77.9%.

They could look at imposing a higher minimum fine to ensure that dog owners keep a close watch on their pets.</td></tr>
</table>

Question No.		Solution
2	(a)	Let μ be <u>mean winnings claim amount.</u> H_0: μ = \$2,000 (or μ ≥ \$2,000) H_1: μ < \$2,000
	(b)	Sample size, n = <u>304</u> Sample mean, $\bar{x}$ = \$1941.94 Sample SD, s = \$440.17
	(c)	P-Value = 0.011 (1-sample t) 95% confidence (Upper) bound for μ is : μ < 1983.59
	(d)	P-Value = 0.011 < 5% Furthermore, the 95% confidence bound does not include hypothesized mean of \$2,000.

Statistics

Variable	N	Mean	StDev
Winnings Claim	304	1941.94	440.17

Test

Null hypothesis H_0: μ = 2000

Alternative hypothesis H_1: μ < 2000

T-Value	P-Value
-2.30	0.011

Descriptive Statistics

N	Mean	StDev	SE Mean	95% Upper Bound for μ
304	1941.9	440.2	25.2	1983.59

μ: population mean of Winnings Claim (\$)

Question No.		Solution

Hence, reject H_0

There is evidence to believe that the average winnings claim is lesser than $2,000

(e)

If the conclusion in (d) is "Reject H_0" (which is the correct answer):
Since we rejected H_0, the conclusion in (d) will not change because if we use a 90% confidence interval/bound, the confidence bound will decrease (or the critical region will increase) to make the analysis to favour H_1 even more.

Descriptive Statistics

N	Mean	StDev	SE Mean	90% Upper Bound for μ
304	1941.94	440.17	25.25	1974.36

μ: population mean of Winnings Claim

3 (a)

Scenario 1:

Condition	Fulfilled?
There is a repeated fixed number of trials or observations, *n*.	**Y** / N
All these trials are independent.	Y / **N**
Each trial should end in one of two outcomes: success or failure.	**Y** / N
The probability of success, *p*, must be the same for all trials.	Y / **N**

Scenario 2:

Condition	Fulfilled?
There is a repeated fixed number of trials or observations, *n*.	**Y** / N
All these trials are independent.	**Y** / N
Each trial should end in one of two outcomes: success or failure.	**Y** / N
The probability of success, *p*, must be the same for all trials.	**Y** / N

(b)(i)

Let X = number of buses that breakdown in a day.

$X \sim B(n = 20, p = 0.03)$

Question No.		Solution
	(b)(ii)	Mean = m = np = 20 x 0.03 = 0.6
	(b)(iii)	P(X = 1) = 0.336
	(b)(iv)	P(X ≤ 1) = P(X = 0) + P(X = 1) = 0.544 + 0.336 =0.88
	(b)(v)	P(X>= 2) = 1 – P(X ≤ 1) = 0.12 Alternative P(X>= 2) = P(X=2) + P(X=3) =0.12. The probabilities for the remaining X values is 0.
	(b)(vi)	Since P(X>= 2) = 0.12. It is not rare to have 2 breakdowns of more per day. So LOCO should hire a full time staff. Since P(X>= 2) = 0.12 is small, LOCO should outsource. Since mean of 0.6, the expected breakdown is 0.6 per day. 0.6 < 2. So LOCO should outsource.
	(c)(i)	$Y \sim N(11, 4)$
	(c)(ii)	P(Y < 10) = 0.3085 = 0.309 (to 3 d.p)
	(c)(iii)	P(10 < Y < 12) = 0.3829 =0.383 (to 3 d.p)
4	(a)(i)	$\bar{T} \sim N\left(30, \frac{64}{n}\right)$ Assumptions: The distribution for journey time per delivery is normally distributed.
	(a)(ii)	P(total journey time > 600) = $P(\mathrm{n}\bar{T} > 600)$ =$P(\bar{T} > {}^{600}/_{16})$ =$P(\bar{T} > 37.5)$ = 0.0000884
	(a)(iii)	P(total journey time > 600) = $P(\mathrm{n}\bar{T} > 600)$

Question No.		Solution
		$=P\left(\bar{T} > {}^{600}/_{17}\right)$ $=P(\bar{T} > 35.294)$ = 0.00406 Yes, since the probability is close to 0 (<5%), it is a rare event.
	(a)(iv)	Maximum delivery is 16 since the probability of the truck driver exceeding the 10 hours driving limit when n is 17 is 0.00406 which is more than 1 in 1000 (or 0.001) whereas when n is 16, the probability is 0.0000884 which is less than 1 in 1000 (or 0.001).
	(b)(i)	Since n = 50 (>30), $\bar{M}$ is approximately normally distributed by <u>Central Limit Theorem</u>. $$\bar{M}\sim N(65, \frac{100}{50} = 2)$$
	(b)(ii)	Using Minitab, $P(\bar{M} > 64)$ $= 0.7603$
	(b)(iii)	If the scripts were from *City XYZ*, we have $\bar{M}\sim N(65{,}2)$ $P(\bar{M} > 68.5) = 0.006657$
	(b)(iv)	Likelihood that we have a sample with mean of 68.5 or more if it is from *City XYZ* is 0.0068 << 0.05, which is <u>rare</u>. Hence we have evidence to conclude that the scripts are <u>not from *City XYZ*</u>.