Math 221: Basic Statistics Exam #1 Solutions

Week #7

(1) In this problem, the study describes the sample chosen from a **larger** population

Population: All dentists
Sample: 3,000 dentists

Population: All cola-drinkers
Sample: 300 consumers

Population: All high schoolers
Sample: 700 students

(2) For this problem, if the answer is a number it's **numerical**, otherwise it's **categorical**.

(a) categorical
(b) numerical
(c) numerical

(a) categorical
(b) numerical
(c) numerical

(a) categorical
(b) numerical
(c) categorical

(3) To construct 5 frequency classes, round up the value from the formula $\frac{Max-Min}{5}$ and then add this amount to the minimum value 5 times. Here's what you get.

$$\frac{261-105}{5} = 31.2$$
$$\text{width} = 32$$

Classes	**Freq.**
105-136	2
137-168	5
169-200	11
201-232	3
233-265	5
Total	26

$$\frac{247-150}{5} = 19.4$$
$$\text{width} = 20$$

Classes	**Freq.**
150-169	7
170-189	2
190-209	8
210-229	3
230-250	6
Total	26

$$\frac{293-141}{5} = 30.4$$
$$\text{width} = 31$$

Classes	**Freq.**
141-171	5
172-202	10
203-233	6
234-264	3
265-296	2
Total	26

(4) Skipping the dotplot, I'll just fill in the table from problem 6

Age(x)	Freq. (f)	xf	x^2f
4	4	16	64
5	2	10	50
6	4	24	144
7	4	28	196
8	4	32	256
9	3	27	243
10	7	70	700
11	5	55	605
12	1	12	144
13	2	26	338
14	4	56	784
Total	40	356	3524

Age(x)	Freq. (f)	xf	x^2f
3	1	3	9
4	3	12	48
5	2	10	50
6	6	36	216
7	2	14	98
8	6	48	384
9	3	27	243
10	4	40	400
11	4	44	484
12	3	36	432
13	2	26	338
14	4	56	784
Total	40	352	3486

Age(x)	Freq. (f)	xf	x^2f
4	5	20	80
5	2	10	50
6	4	24	144
7	3	21	147
8	2	16	128
9	6	54	486
10	5	50	500
11	6	66	726
12	3	36	432
13	3	39	507
14	1	14	196
Total	40	350	3396

(5) Using the previous bowling data, we compute 4 of the 5 M's. The midrange is the average of the highest and lowest values. The mode(s) is the most frequent value. Since our data set has 26 values, the median is the average of the 13th and 14th value. The mean is the sum divided by 26.

$$\begin{aligned} \text{Midrange} &= \frac{261+105}{2} = 183 \\ \text{Mode(s)}: &= 167, 181, 198 \\ \text{Median}(\tilde{\text{X}}) &= \frac{182+186}{2} = 184 \\ \text{Mean}(\overline{\text{X}}) &= \frac{4949}{26} \approx 190.346 \end{aligned}$$

$$\begin{aligned} \text{Midrange} &= \frac{247+150}{2} = 198.5 \\ \text{Mode(s)}: &= 242 \\ \text{Median}(\tilde{\text{X}}) &= \frac{196+197}{2} = 196.5 \\ \text{Mean}(\overline{\text{X}}) &= \frac{5167}{26} \approx 198.731 \end{aligned}$$

$$\begin{aligned} \text{Midrange} &= \frac{293+141}{2} = 217 \\ \text{Mode(s)}: &= 178, 188 \\ \text{Median}(\tilde{\text{X}}) &= \frac{188+189}{2} = 189.5 \\ \text{Mean}(\overline{\text{X}}) &= \frac{5187}{26} = 199.5 \end{aligned}$$

(6) Using the charts for housefly data, we can compute the range, variance and standard deviation.

$$\begin{aligned} \text{Range} &= 14-4=10 \\ \text{Variance}(\text{S}^2) &= \frac{3524-\frac{(356)^2}{40}}{40-1} \\ &\approx 9.118 \\ \text{Std.Dev.(S)} &= \sqrt{\text{Variance}} \\ &\approx 3.020. \end{aligned}$$

$$\begin{aligned} \text{Range} &= 14-3=11 \\ \text{Variance}(\text{S}^2) &= \frac{3486-\frac{(352)^2}{40}}{40-1} \\ &\approx 9.959 \\ \text{Std.Dev.(S)} &= \sqrt{\text{Variance}} \\ &\approx 3.156. \end{aligned}$$

$$\begin{aligned} \text{Range} &= 14-4=10 \\ \text{Variance}(\text{S}^2) &= \frac{3396-\frac{(350)^2}{40}}{40-1} \\ &\approx 8.551 \\ \text{Std.Dev.(S)} &= \sqrt{\text{Variance}} \\ &\approx 2.924. \end{aligned}$$

(7) To solve this problem, the first thing that you must do is sort your ten numbers in numeric order from smallest to largest. Then, the five number summary requires you to find the following five numbers: Min, Q1, Median, Q3 and Max. Of ten numbers, the Q1 is the 3rd number and Q3 is the 8th number in the list. The median is the average of the 5th and 66th number in the list. The interquartile value is just Q3-Q1, while the midhinge is just $\frac{Q3+Q1}{2}$. A box-and-whisker plot just represents your five-number summary graphically on a number line.

$$\begin{aligned} \text{Maximum} &= 91 \\ \text{Q}_3 &= 53 \\ \text{Median} &= \frac{43+46}{2} = 44.5 \\ \text{Q}_1 &= 29 \\ \text{Minimum} &= 20 \\ \text{Interquartile} &= 53-29=24 \\ \text{Midhinge} &= \frac{53+29}{2} = 41 \end{aligned}$$

$$\begin{aligned} \text{Maximum} &= 97 \\ \text{Q}_3 &= 49 \\ \text{Median} &= \frac{35+35}{2} = 35 \\ \text{Q}_1 &= 30 \\ \text{Minimum} &= 22 \\ \text{Interquartile} &= 49-30=19 \\ \text{Midhinge} &= \frac{49+30}{2} = 39.5 \end{aligned}$$

$$\begin{aligned} \text{Maximum} &= 88 \\ \text{Q}_3 &= 47 \\ \text{Median} &= \frac{39+43}{2} = 41 \\ \text{Q}_1 &= 26 \\ \text{Minimum} &= 22 \\ \text{Interquartile} &= 47-26=21 \\ \text{Midhinge} &= \frac{47+26}{2} = 36.5 \end{aligned}$$

(8) To solve this problem, just note that failing is the opposite of passing. Also, if you pass the exam, you may either receive an 80 and above or score below an 80.

$$\begin{aligned} P(\text{Fail}) &= 100\% - 64\% \\ &= 36\% \\ P(\text{Pass} < 80) &= 64\% - 27\% \\ &= 37\% \end{aligned}$$

$$\begin{aligned} P(\text{Fail}) &= 100\% - 62\% \\ &= 38\% \\ P(\text{Pass} < 80) &= 62\% - 29\% \\ &= 33\% \end{aligned}$$

$$\begin{aligned} P(\text{Fail}) &= 100\% - 67\% \\ &= 33\% \\ P(\text{Pass} < 80) &= 67\% - 24\% \\ &= 43\% \end{aligned}$$

(9) This problem requires a bit of conditional probability. To pick two items from a box of a single type **replacement**, you must remember that after selecting the first item, there are fewer remaining items in the box. Also, the opposite of "at least one part is defective" is "both parts are not defective."

$$\begin{aligned} P(\text{2 BAD}) &= \frac{3}{7}\cdot\frac{2}{6} \\ &= \frac{1}{7} \approx 14.3\% \\ P(\text{2 OK}) &= \frac{4}{7}\cdot\frac{3}{6} \\ &= \frac{2}{7} \approx 28.6\% \\ P(\text{1+ BAD}) &= 1 - P(\text{2 OK}) \\ &= 1 - \frac{2}{7} = \frac{5}{7} \approx 71.4\% \end{aligned}$$

$$\begin{aligned} P(\text{2 BAD}) &= \frac{4}{15}\cdot\frac{3}{14} \\ &= \frac{2}{35} \approx 5.71\% \\ P(\text{2 OK}) &= \frac{11}{15}\cdot\frac{10}{14} \\ &= \frac{11}{21} \approx 52.4\% \\ P(\text{1+ BAD}) &= 1 - P(\text{2 OK}) \\ &= 1 - \frac{11}{21} = \frac{10}{21} \approx 47.6\% \end{aligned}$$

$$\begin{aligned} P(\text{2 BAD}) &= \frac{5}{12}\cdot\frac{4}{11} \\ &= \frac{5}{33} \approx 15.2\% \\ P(\text{2 OK}) &= \frac{7}{12}\cdot\frac{6}{11} \\ &= \frac{7}{22} \approx 31.8\% \\ P(\text{1+ BAD}) &= 1 - P(\text{2 OK}) \\ &= 1 - \frac{7}{22} = \frac{15}{22} \approx 68.2\% \end{aligned}$$

(10) Filling in the table for this problem gives

	In Time		
Satisfied	*Yes*	*No*	*Total*
Yes	1197	33	1230
No	127	143	270
Total	1324	176	1500

	In Time		
Satisfied	*Yes*	*No*	*Total*
Yes	1117	53	1170
No	147	183	330
Total	1264	236	1500

	In Time		
Satisfied	*Yes*	*No*	*Total*
Yes	1037	73	1110
No	167	223	390
Total	1204	296	1500

Now, if S represents being "satisfied" and R represents "receiving the product in time," then

$$P(\text{S and not R}) = \frac{33}{1500} \approx 2.2\%$$
$$P(\text{S or R}) = \frac{1324 + 1230 - 1197}{1500}$$
$$= \frac{1357}{1500} \approx 90.4\overline{66}\%$$

$$P(\text{S and not R}) = \frac{53}{1500} \approx 3.5\overline{33}\%$$
$$P(\text{S or R}) = \frac{1264 + 1170 - 1117}{1500}$$
$$= \frac{1317}{1500} \approx 87.1\overline{33}\%$$

$$P(\text{S and not R}) = \frac{73}{1500} \approx 4.8\overline{66}\%$$
$$P(\text{S or R}) = \frac{1204 + 1110 - 1037}{1500}$$
$$= \frac{1277}{1500} \approx 87.8\%$$

[EC] To solve the extra credit problem, sort the given 15 items. The, the 15th percentile value is the $15\left(\frac{15}{100}\right) = 2.25 \rightarrow$3rd number in the list which is **9.72**. The 90th percentile value is the $15\left(\frac{90}{100}\right) = 13.5 \rightarrow$14th value which is **60.8**.