MG221: Applied Probability & Statistics
Final 2019

1. The Air Quality Index (AQI), reported as a 24 hour average value at a monitoring station located on Lodhi Road, New Delhi for the 30 days in November 2019, in sorted order, were as follows:

91, 126, 130, 163, 163, 174, 178, 179, 182, 184, 193, 196, 199, 212, 227
251, 252, 252, 254, 263, 287, 313, 328, 350, 398, 424, 434, 436, 458, 553

World Health Organization (WHO) classifies an AQI level between 200 and 300 as Very Unhealthy and above 300 as Hazardous. Given that the mean and standard deviation of the above 30 AQI levels to be 261.6667 and 114.3426 respectively, answer the following:

a. Is there sufficient evidence to say that the average AQI in New Delhi in November, 2019 was at a Very Unhealthy to a Hazardous level? **[4]**

b. Give a 95% confidence interval for the average AQI level in New Delhi in November, 2019 and interpret it in terms of the WHO classification. **[3+1=4]**

c. Is there sufficient evidence to say that the AQI in New Delhi was at a Hazardous level more than 25% of the time in November, 2019? **[4]**

2. One of the questions in a survey of working professionals in the UK was to order fourteen factors like Hard-Work, Intelligence, Grit, Looks, Employer Support, Luck etc. according to their importance in professional career progression for a woman. The same question was also asked about the professional career progression of a man. In this problem, the interest lies in the difference in perceptions about Hard-Work as the most important factor for professional career progression of the two genders, and this response's interaction with the gender of the respondents themselves. Out of the 100 respondents, 40 were women, 30 of whom cited Hard-Work as the most important factor for professional career progression of a woman, as did as many men. A total of 47 respondents cited Hard-Work as the most important factor for professional career progression of a man, 35 of them being men themselves. 20 men responded Hard-Work as the most important factor for both men and women professionals, while 23 respondents did not cite Hard-Work as the most important factor for either the men or the women professionals. Answer the following:

a. Is there a significant difference in the men and women professionals' perceptions in Hard-Work as the most important factor for professional career progression of a woman? **[4]**

b. Is there a significant difference in the men and women professionals' perceptions in Hard-Work as the most important factor for professional career progression of a man? **[4]**

c. Do the men differ in their perceptions about Hard-Work as the most important factor for career progression for a man vis-à-vis a woman professional? **[6]**

d. Similarly do the women also differ in their perceptions about Hard-Work as the most important factor for career progression for a man vis-à-vis a woman professional? **[6]**

e. Based on your findings in **a** through **d** above, briefly summarise how the men and women professionals differ in their perceptions about Hard-Work as the most important factor for professional career progression of a woman compared to that of a man. **[5]**

3. A Blue-Light-Sale (BLS) with respect to retail stores, refers to the act of hoisting a big, tall, rotating blue light in the middle of the store and selling everything in the store at a discount price, while the blue-light is on. A BLS was administered in 8 stores of a garment-accessory-furniture-artifact-selling retail chain in Bengaluru, during 7:00-8:00 PM on a Saturday evening. The sales (in nearest thousands of Rs.), in the hour just preceding the BLS (6:00-7:00 PM), and the hour of the BLS (7:00-8:00 PM) in these 8 stores are observed as follows:

Store	1	2	3	4	5	6	7	8
6:00-7:00 PM	112	86	65	150	45	70	123	148
7:00-8:00 PM	134	101	83	184	64	88	156	176

Answer the following:

a. Is there sufficient evidence to conclude that the expected sale during the BLS-hour was 20% more than the expected sale in the hour just preceding it? Assume Normality but you must clearly write-down the expression of the random variable, which is assumed to be Normal. **[8+2=10]**

b. Answer the same as in **a** without the assumption of Normality. **[7]**

c. Can the (statistically significant) increase in expected sales during the BLS-hour be attributed to the causal effect of the BLS? Explain why or why not. **[3]**

4. In a study of long-term effect of pre-mature birth and being fed formula-milk as an infant on cardio-vascular health, doctors measured the combined volume of the two ventricles in healthy 23-28 year olds using MRI scans. Some of these individuals were born on time, while the others were pre-maturely born. Some had mainly nursed on their mothers' milk, while the others were fed only formula-milk during their infancy. The number of individuals in each of these categories and the mean and standard deviations of their ventricles volume (in ml) are as follows:

Ventricles Volume	Born on Time	Pre-Maturely Born
Mother's Milk	$n = 30$, Mean=171.4, SD=5.6848	$n = 12$, Mean=166.5, SD=6.9348
Formula-Milk	$n = 20$, Mean=168.8, SD=5.8183	$n = 8$, Mean=158.75, SD=8.5482

Assume the Ventricles Volumes to be Normally distributed in each group. Answer the following:

a. Show that the pre-maturely born have smaller ventricles volume compared to those born on time. **[10]**

b. Show that though there is no significant difference in ventricles volume between the Mother's-Milk and Formula-Milk groups for those born on time, for the pre-maturely born, mother's milk make a significant difference in ventricles volume over formula-milk. **[5+5=10]**

c. Strengthen the last conclusion using an appropriate distribution-free test, given the ventricles volumes of the 20 pre-maturely born individuals categorised according to the feed of their infancy are as follows: **[8]**

Mother's Milk:	150	158	164	164	166	166	170	170	172	172	172	174
Formula-Milk:	148	150	154	156	158	164	168	172				

5. In a study of the association of Entrepreneurial Orientation (EO - measured in terms of innovativeness, risk-taking, proactiveness, competitive aggressiveness and autonomy of its employees) with the Profitability of firms, a random sample of 450 firms are cross-classified according to their EO being low, moderate or high and Profitability being either low or high in the following contingency table:

EO $\rightarrow$ Profitability $\downarrow$	Low	Moderate	High
Low	25	120	155
High	5	45	100

Answer the following:

a. Based on the above data, will it be proper to conclude the existence of a significant association between the EO-levels and Profitability of the firms? **[7]**

b. Briefly describe the nature of this association (using only the above data). **[3]**

6. Consider a negative binomial sampling in which one keeps on performing i.i.d. Bernoulli trials till there are n Successes. Let Y denote the number of Failures with $Y \sim \text{NB}(n, \pi)$, where π=P(Success). Answer the following:

a. Show that the MLE of π is $\frac{n}{Y+n}$ but its UMVUE is $\frac{n-1}{Y+n-1}$. **[3+7=10]**

b. If in order to get a total of 30 Heads in a sequence of independent tosses, the total number of Tails turns out to be 20, is there sufficient evidence to conclude that the coin being tossed is biased? **[5]**