

MG221: Applied Probability & Statistics

Final 2018

1. Annual returns, in percentages, on NAV (Net Asset Value) of 48 different equity linked mutual funds (ELMFs), managed by a certain financial institution, in a certain year (in sorted order) are as follows:

0.29	1.21	1.56	1.83	3.31	3.41	3.42	3.60	3.92	4.05	4.11	4.24
4.35	4.41	4.48	4.75	5.03	5.17	5.20	5.22	5.40	5.60	5.71	5.85
6.10	6.12	6.18	6.88	6.90	7.22	7.42	7.62	7.63	7.67	7.87	7.99
8.06	8.21	8.46	8.60	8.65	8.87	9.64	10.01	10.74	11.33	11.57	12.52

It is given that the mean and standard deviation of the above returns are 6.2162 and 2.7514 respectively, and that year the inflation rate was 5% and the capital gains tax rate was 10%. Real return is the return in excess of this inflation and capital gains tax (for example, the real return for the return figure of 5.40% above is $5.40 - (0.1 \times 5.40) - 5 = -0.14\%$). Treating the above as a representative sample of ELMFs managed by that financial institution, answer the following:

- a. Is there sufficient evidence to expect the real returns of the ELMFs managed by that financial institution to be positive? [4]
- b. Give 95% confidence intervals for the mean and variance of the real returns of the ELMFs managed by that financial institution. What assumptions are required for the validity of these intervals? [4+4+2=10]
- c. Is there sufficient evidence for concluding that the first quartile of the real returns of the ELMFs managed by that financial institution is negative? [4]
- d. What can be said about the minimum value of the probability of the real return of an ELMF managed by that financial institution being negative, with approximate 99% confidence? [4]
- e. Will investing in an ELMF managed by that financial institution be a good decision? Based on your analyses in parts a through d above, explain why or why not. [2]

2. A study was undertaken to see whether food package designs signaling health versus hedonism have any association with the healthiness of the food contained in the package, as perceived by the consumers. The food product chosen for the study was yogurt. The two designs of the packages were chosen following empirical findings in the literature regarding perceived healthiness and hedonism of a food-package. For instance, the healthy package design used only blue, green and white colours; contained text focusing on health and nutrition information like calories, fat, sugar and salt content per serving; and had the brand name pointing downwards. The hedonistic package design on the other hand used only red, orange and brown colours; contained text focusing on flavour and taste of the yogurt; and had the brand name pointing upwards. Though the packagings were different, both contained the same yogurt of the promised taste, flavour and nutrition content. Since response to the package design is known to be affected by the attitude of the consumers, they are first segmented into price-sensitive versus health-conscious categories. A discount super-market and an expensive organic-only green-market are chosen to obtain responses from the two respective customer segments. In both the stores, 100 customers were randomly chosen after their check-out with 50 of them randomly being offered the healthy looking package while the other 50 being offered the hedonistic looking package. The packages were offered free of cost, in exchange of a simple Yes or No response to the question, "Did the yogurt taste healthy?" 58 super-market customers answered Yes, with 38 of them having tasted it from the healthy looking package, while 70 green-market customers answered No, with 36 of them having tasted it from the hedonistic looking package. Answer the following:

- a. Show that if one ignores customer segmentation in this data set, then as such consumers are not more likely to perceive the taste of yogurt in the healthy looking package to be healthy, but they definitely perceive the taste of yogurt in the hedonistic looking package to be less healthy. [4+4=8]
- b. However, show that while the green-market customers' perception of healthy taste has no significant association with the package design, super-market customers on the other hand are more likely to perceive yogurt in healthy looking package to also taste healthier. [4+4=8]

3. In a Université de Liège (Belgium) study on the association of nicotine with vigilance, published in *Psychopharmacology* (1999), 16 male volunteers, all of whom were between 18 and 25 years old and smoked at least 15 cigarettes a day, were first randomly divided into two groups of 8 each. A matrix-type patch containing 21 mg S(-)-nicotine was used to transdermally deliver nicotine to the volunteers in one group, while a placebo patch was attached to the volunteers in the other group. Volunteers in both the groups abstained from smoking for at least 10 hours prior to

participating in the study. Vigilance was assessed using the Rapid Visual Information Processing Test (RVIPT). In RVIPT, a series of digits, generated by a computer, is presented on a visual display unit at the rate of 100 digits per minute, and the test takers are instructed to press a response button as quickly as possible when they detect sequences of 3 consecutive odd or 3 consecutive even digits. This is called a hit. Any two sequences or hits were separated by a minimum of 5 and a maximum of 30 digits. An average of 120 hits were presented in a 15-minute RVIPT. Volunteers from both the groups took the 15-minute RVIPT twice, once just before the placebo/nicotine patch was attached to them, and then once again one and half hours after the attachment of the patch. The number of correct hits thus scored by the 16 volunteers in the two RVIPTs are as follows:

Placebo	Vol. No.	1	2	3	4	5	6	7	8
Patch	Before	48	53	81	61	83	58	75	60
Group	After	49	45	82	51	84	64	67	70
Nicotine	Vol. No.	1	2	3	4	5	6	7	8
Patch	Before	74	58	45	70	66	57	61	85
Group	After	86	65	53	81	69	58	58	98

Answer the following:

- Perform a distribution-free test to ensure that the two groups are homogeneous with respect to their RVIPT scores before the attachment of the patches. Why is this confirmation necessary for the study? [8+2=10]
- Perform an appropriate t -test to check whether there is any significant difference between the before and after RVIPT scores in the placebo patch group. [6]
- Perform a distribution-free test to check whether there is any significant improvement in the RVIPT scores one and half hours after the attachment of the nicotine patches. [6]
- Perform an appropriate t -test to check whether the improvement in RVIPT scores one and half hours after the attachment of the nicotine patches is significantly better compared to the placebo patches. [8]
- Write a brief report on the association of nicotine with vigilance, based on your analyses of the data carried out in parts **a**, **b**, **c** and **d** above. [3]
- Comment on the aptness of using only smokers as study subjects in the experiment. [2]
- Is the placebo group really necessary in this experiment? Explain why or why not. [2]
- In lights of your answers to parts **f** and **g** above, how could you have better designed the above experiment? [3]

4. A random sample of 25 four-wheel, passenger-car, traffic violation records in a city is cross-classified according to the type of drivers, and nature of traffic violations. Drivers are classified as either owners or chauffeurs, while traffic violations are categorised as parking violations, speeding violations or traffic light violations. The frequencies of their occurrences in the sample of 25 are as follows:

Violation Type → Driver Type ↓	Parking	T.Lights	Speeding
Owner	2	3	5
Chauffeur	7	4	4

Answer the following:

- Is there any evidence to say that the nature of traffic violations committed by the the owner-drivers is significantly different from that of the chauffeur-drivers? [5]
- Is there any significant evidence to say that the chauffeur-drivers are more likely to commit a parking violation than owner-drivers? [5]

5. Answer the following:

- Show that the size- α fixed significance level Likelihood Ratio Test for testing the null hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against the alternative $H_a : \sigma^2 > \sigma_0^2$, based on an i.i.d. sample $Y_1, Y_2, \dots, Y_n \sim N(\mu, \sigma^2)$ is given by, “Reject H_0 if $\sum (Y_i - \bar{Y})^2 > \chi_{1-\alpha, n-1}^2$ ”, where $\chi_{1-\alpha, n-1}^2$ is the $(1 - \alpha)$ -th quantile of χ_{n-1}^2 . [8]
- Let $Y_{11}, Y_{12}, \dots, Y_{1n_1}$ i.i.d. $N(\mu_1, \sigma^2)$ and $Y_{21}, Y_{22}, \dots, Y_{2n_2}$ i.i.d. $N(\mu_2, \sigma^2)$ be independent. Show that the p -value for testing the null hypothesis $H_0 : \mu_1 - \mu_2 = \mu_0$ against the alternative $H_a : \mu_1 - \mu_2 \neq \mu_0$ based on the Likelihood Ratio Test is given by $2P(t_{n_1+n_2-2} > |t_{obs}|)$, where $t_{obs} = (\bar{Y}_1 - \bar{Y}_2 - \mu_0) / \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$, for $i = 1, 2$, $\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$ and $s_p^2 = \frac{1}{n_1+n_2-2} \sum_{i=1}^2 \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$. [12]