NOT TO BE TAKEN AWAY

THE UNIVERSITY OF HONG KONG
DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

STAT2602 PROBABILITY AND STATISTICS II

December 9, 2020 Time: 9:30 a.m. - 11:30 a.m.

Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is candidates' responsibility to ensure that their calculator operates satisfactorily, and candidates must record the name and type of the calculator used on the front page of the examination script.

Answer ALL SEVEN questions. Marks are shown in square brackets.

1. Suppose that two independent random variables X and Y have the density functions $f_X(x)$ and $f_Y(x)$, respectively, where

$$f_X(x) = \begin{cases} 4e^{-4x}, & \text{for } x > 0; \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad f_Y(x) = \begin{cases} 6e^{-6x}, & \text{for } x > 0; \\ 0, & \text{otherwise.} \end{cases}$$

   (i) Calculate the moment generating function of X. **[4 marks]**

   (ii) Calculate the values of $\mathrm{E}[X]$ and $\mathrm{E}[X^2]$. **[4 marks]**

   (iii) Calculate the moment generating function of $X + Y$. **[4 marks]**

   (iv) Calculate the moment generating function of $X - Y$. **[4 marks]**

   (v) Does there exist a constant $c > 0$ such that cY and X have the same distribution? Explain your reason. **[4 marks]**

[Total: 20 marks]

2. Consider an independent random sample $\mathbf{X} = \{X_1, X_2, ..., X_{2n}\}$ from $\mathrm{N}(1, 1)$. Based on $\mathbf{X}$, define another two random samples

$$\mathbf{Y} = \{Y_1, Y_2, ..., Y_n\} \quad \text{and} \quad \mathbf{Z} = \{Z_1, Z_2, ..., Z_n\},$$

where

$$Y_i = X_{2i-1} - X_{2i} \quad \text{and} \quad Z_i = \mathrm{I}(Y_i > 0) - \mathrm{I}(Y_i < 0)$$

for $i = 1, 2, ..., n$. Here, $\mathrm{I}(\cdot)$ is the indicator function.

   (i) Calculate $\mathrm{E}(Y_1)$ and $\mathrm{Var}(Y_1)$. **[4 marks]**

   (ii) For $n = 50$, approximate $\mathrm{P}\Big(-\frac{1}{4} < \overline{Y} < \frac{1}{4}\Big)$, where $\overline{Y}$ is the sample mean of $\mathbf{Y}$. **[5 marks]**

   (iii) For $n = 50$, approximate $\mathrm{P}\Big(-\frac{1}{4} < \overline{Z} < \frac{1}{8}\Big)$, where $\overline{Z}$ is the sample mean of $\mathbf{Z}$. **[5 marks]**

   (iv) For a large value of n, approximate $\dfrac{\mathrm{P}(|\overline{Z}| < \frac{1}{4})}{2 + \mathrm{P}(|\overline{Z}| > \frac{1}{2})}$. **[6 marks]**

[Total: 20 marks]

3. Let $\{X_1, X_2, ..., X_n\}$ be an independent random sample from $\mathrm{N}(1, \theta)$, where $\theta > 0$.

   (i) Find the MLE of θ on the parameter space $\Omega = \{\theta : 1 \le \theta \le 2\}$. **[10 marks]**

   (ii) Write down the formulas of bias and MSE of the MLE in (i) by using chi-squared distribution and density functions. **[10 marks]**

[Total: 20 marks]

4. Let $\{X_1, X_2, ..., X_n\}$ be an independent positive random sample from a population with density function given by

$$f(x;\theta) = c_\theta x^{\frac{\alpha}{2}} \exp(-\beta x) \quad \text{for } x > 0,$$

where the unknown parameter vector $\theta = (\alpha, \beta)$ with $\alpha \in \mathcal{R}$ and $\beta > 0$, and $c_\theta > 0$ is a constant only depending on θ.

(i) Find a sufficient and complete statistic for θ. **[4 marks]**

(ii) If $\mathrm{E}(X_1) = 2\beta$ and $\mathrm{E}(\log X_1) = 2\alpha + \beta$, find the UMVUE of θ. **[6 marks]**

[Total: 10 marks]

5. Let X_1, X_2, X_3, X_4, X_5 be an independent random sample from $\mathrm{N}(\mu, \sigma^2)$, and the following data are observed:

$$x_1 = 50, \quad x_2 = 65, \quad x_3 = 40, \quad x_4 = 40, \quad x_5 = 60.$$

(i) Find a 90% confidence interval estimate of μ. **[5 marks]**

(ii) Find a 90% confidence interval estimate of σ^2. **[5 marks]**

[Total: 10 marks]

6. Suppose 90 randomly selected HKU students are surveyed to determine if they have the preference to play basketball, football, and baseball. Of the 90 surveyed, 35 reported preference to play basketball, 25 reported preference to play football, and 30 reported preference to play baseball. Can we conclude that HKU students have no preference to play these three kinds of sports? Test at the 10% significance level.
[Total: 10 marks]

7. Let $\mathbf{X} = \{X_1, X_2, ..., X_n\}$ be an independent positive random sample from a population with the density function given by

$$f(x;\theta) = \begin{cases} \theta e^{-\theta x}, & \text{for } x > 0; \\ 0, & \text{otherwise.} \end{cases}$$

(i) Find the rejection region of the most powerful test for hypotheses:

$$H_0 : \theta = 4 \quad \text{versus} \quad H_1 : \theta = 2.$$

[5 marks]

(ii) Find the critical value such that this test has an asymptotic size 0.05 when n is large. **[5 marks]**

[Total: 10 marks]

A LIST OF STATISTICAL FORMULAE

1. $M_X(t) = \mathrm{E}(e^{tX})$. $\quad \mathrm{E}(X^r) = \left(\dfrac{\mathrm{d}^r}{\mathrm{d}t^r} M_X(t) \right) \Big|_{t=0}$.

2. $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$. $\quad S^2 = \dfrac{1}{n}\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$.

3. LLN: $\overline{X} \to_p \mu$ for large n.

4. CLT: $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$ for large n.

5. Normal population $\Longrightarrow \dfrac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$, $\dfrac{\overline{X} - \mu}{S/\sqrt{n-1}} \sim t_{n-1}$.

6. $\mathrm{Bias}(\hat{\theta}) = \mathrm{E}(\hat{\theta}) - \theta$. $\quad \mathrm{E}\left[(\hat{\theta} - \theta)^2\right] = \mathrm{Var}(\hat{\theta}) + \left[\mathrm{Bias}(\hat{\theta})\right]^2$.

7. $I(\theta) = \mathrm{E}\left[\left(\dfrac{\partial \ln f(X;\theta)}{\partial \theta} \right)^2 \right] = \mathrm{E}\left[-\dfrac{\partial^2 \ln f(X;\theta)}{\partial \theta^2} \right]$. $\quad \mathrm{Var}(\hat{\theta}) \geq \dfrac{1}{nI(\theta)}$.

8. $\mathbf{f}(x_1, x_2, \ldots, x_n; \theta) = g\left(u(x_1, x_2, \ldots, x_n), \theta\right) \cdot h(x_1, x_2, \ldots, x_n)$.

9. Expontial family: $f(x;\theta) = h(x)c(\theta)\exp\left(\sum_{i=1}^{s} p_i(\theta) t_i(x)\right)$.

10. $\overline{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$. $\quad \overline{x} \pm t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n-1}}$. $\quad \left(\dfrac{ns^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{ns^2}{\chi^2_{1-\alpha/2,n-1}} \right)$.

11. $\overline{x} - \overline{y} \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_x^2}{n_1} + \dfrac{\sigma_y^2}{n_2}}$. $\quad \overline{x} - \overline{y} \pm t_{\alpha/2,n_x+n_y-2}\sqrt{\dfrac{n_x s_x^2 + n_y s_y^2}{n_x + n_y - 2}\left(\dfrac{1}{n_x} + \dfrac{1}{n_y}\right)}$.

12. $-2\ln\Lambda \approx \sum_i \dfrac{(O_i - E_i)^2}{E_i} = \sum_i \dfrac{O_i^2}{E_i} - n$. $\quad -2\ln\Lambda \approx n\left(\sum_i \sum_j \dfrac{O_{i,j}^2}{n_{i.}n_{j.}} - 1 \right)$.

Cumulative distribution function $\Phi(x)$ for $N(0,1)$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

αth upper percentile for the student's t distribution t_ν

	α					
ν	10.0%	5.0%	2.5%	1.0%	0.5%	0.1%
1	3.078	6.314	12.706	31.821	63.657	318.31
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
∞	1.282	1.645	1.960	2.326	2.576	3.090

αth upper percentile for the chi-squared distribution χ^2_ν

ν	99.5%	99.0%	97.5%	95.0%	90.0%	10.0%	5.0%	2.5%	1.0%	0.5%
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

********** END OF PAPER **********