NOT TO BE TAKEN AWAY

THE UNIVERSITY OF HONG KONG
DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

STAT2602 Probability and statistics II

20 December, 2024 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 FIVE questions. Marks are shown in square brackets.

1. Management training programs are often instituted to teach supervisory skills and thereby increase productivity. Suppose a company psychologist administers a set of examinations to each of 10 supervisors before such a training program begins and then administers similar examinations at the end of the program. The examinations are designed to measure supervisory skills, with higher scores indicating increased skill. Denote $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$. The results of the tests are shown below:

Supervisor	Pre-Test	Post-Test	Difference
1	63	78	15
2	93	92	-1
3	84	91	7
4	72	80	8
5	65	69	4
6	72	85	13
7	91	99	8
8	84	82	-2
9	71	81	10
10	80	87	7
Sum	775	844	69
S^2	98.25	63.64	26.49

Assume that the pre-test score, post-test score, and the difference of post-test score and pre-test score, all follow normal distributions. We wish to test whether the data provide evidence that the training program is effective in increasing supervisory skills, as measured by the examination scores.

(a) State the appropriate null and alternative hypotheses.

[5 marks]

(b) Calculate the t test statistic.

[5 marks]

(c) Determine whether you would reject H_0 at the 5% significance level. State the reason.

[5 marks]

[Total: 15 marks]

2. Given a random sample $\{X_1, X_2, \ldots, X_n\}$ of size n from the gamma distribution with a probability density function

$$f(x|\theta, k) = \frac{x^{k-1}e^{-x/\theta}}{\Gamma(k)\theta^k}, \quad x \geq 0$$

where θ is the unknown parameter and k is known.

(a) Find the maximum likelihood estimator $\hat{\theta}$ for θ.

[5 marks]

(b) Show that $\hat{\theta}$ is an unbiased estimator of θ. (Hint $\int_0^\infty u^k e^{-u}\,du = \Gamma(k+1)$ and $\Gamma(k+1) = k\Gamma(k)$).

[5 marks]

(c) Show that $\hat{\theta}$ is the unique UMVUE of θ.

[5 marks]

(d) Describe the asymptotic distribution of $\hat{\theta}$.

[5 marks]

[Total: 20 marks]

3. Let $X_1, X_2, \cdots, X_n$ be i.i.d random variables, each following a Uniform distribution $U(0, \theta)$, for some unknown parameter $\theta > 0$.

(a) Find the maximum likelihood estimator of θ and write down the cumulative distribution function of the MLE.

[5 marks]

(b) Based on cumulative distribution function obtained in (a), construct a pivotal quantity involving θ.

[5 marks]

(c) Deduce a $(1-\alpha)\%$ equal tailed confidence interval for θ.

[5 marks]

(d) Illustrate your answer to (c) with $\alpha = 0.05$ and the dataset $X_1 = 0.5$, $X_2 = 0.7$, $X_3 = 1.0$, $X_4 = 1.2$, $X_5 = 1.5$, $X_6 = 1.8$.

[5 marks]

(e) Find the UMVUE of θ^2.

[5 marks]

[Total: 25 marks]

4. Let $X_1, X_2, \ldots, X_{15}$ denote a random sample from the density function

$$f(x \mid \theta) = \frac{1}{\theta} 4x^3 e^{-x^4/\theta} \text{ for } x > 0$$

where $\theta > 0$ is an unknown parameter. (Hint: A nice property of this density is that $2X_i^4/\theta \sim \chi_2^2$, the chi-square distribution with degree of freedom 2.)

(a) We are interested to test $H_0 : \theta = 2$ versus $H_1 : \theta = \theta_1$ with $\theta_1 > 2$. Construct the rejection region for the most powerful test at significance level $\alpha = 0.05$. (Note: Please write down the rejection region explicitly.)

[6 marks]

(b) If you observe $\sum_{i=1}^{15} x_i^4 = 46.98$, what is the p-value?

[5 marks]

(c) Would you reject H_0 at level $\alpha = 0.05$ based on part (b)?

[4 marks]

(d) What is the approximate power of your most powerful test at $\theta_1 = 5$?

[5 marks]

[Total: 20 marks]

ν	Lower percentile for the chi-square distribution χ_ν								
	0.034	0.05	0.1	0.25	0.75	0.9	0.95	0.975	0.99
15	6.68	7.26	8.55	11.04	18.24	22.31	25.00	27.49	30.58
30	17.51	18.49	20.60	24.48	34.80	40.26	43.77	46.98	50.89
40	25.31	26.51	29.05	33.66	45.61	51.81	55.76	59.34	63.69

Note: To read the table, for example, $P(Y < 6.68) = 0.034$, where Y follows a chi-square distribution with degree of freedom 15.

5. Let $X_1, X_2, \cdots, X_n$ denote a random sample from the density function

$$f(x) = \begin{cases} \exp\{-(x-\mu)\}, & x \geq \mu; \\ 0, & \text{otherwise}, \end{cases}$$

where μ is an unknown parameter.

(a) To test $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$, show that the generalized likelihood ratio test takes the form of rejecting H_0 when

$$X_{(1)} \geq c,$$

where $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and c is certain constant.

[5 marks]

(b) From now on suppose that $\mu_0 = 1$, so we aim to test $H_0 : \mu \leq 1$ versus $H_1 : \mu > 1$.

(i) Find the power function of the test in (a).

[5 marks]

(ii) What choice of c would make the test in (a) have size equal to α?

[5 marks]

(iii) How large should n be so that the test in (a) has power $1 - \beta$ for $\mu = 2$ given the value of c you found in (b)(ii)?

[5 marks]

[Total: 20 marks]

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

A LIST OF STATISTICAL FORMULAE

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

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

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

4. Uniform distribution $U(a,b)$ pdf: $f(x) = \frac{1}{b-a}$ for $x \in [a,b]$.

5. Chi-square distribution χ_k^2's pdf: $f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$ for $x > 0$.

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

7. $\chi^2_\nu \sim \sum_{i=1}^{\nu} Z_i^2$. $\quad t_\nu \sim \frac{Z}{\sqrt{\chi^2_\nu/\nu}}$. $\quad F_{\nu_1,\nu_2} \sim \frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}$.

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

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

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

11. $\frac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}$. $\quad \frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$.

12. $\frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$. $\quad \frac{\overline{X}-\overline{Y}-\delta}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$, where $S_p^2 = \frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}$.

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

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

15. $F_{1-\alpha,m,n} = \frac{1}{F_{\alpha,n,m}}$.

16. $-2\ln\Lambda \to \chi^2_d$ as $n \to \infty$.

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

Upper percentile for the student's t distribution t_ν

ν	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090