Student Number

Queen's University
Department of Mathematics and Statistics

MTHE/STAT 353

Final Examination April 24, 2018
Instructor: G. Takahara

- PLEASE NOTE: "Proctors are unable to respond to queries about the interpretation of exam questions. Do your best to answer exam questions as written."
- "The candidate is urged to submit with the answer paper a clear statement of any assumptions made if doubt exists as to the interpretation of any question that requires a written answer."
-
- Formulas and tables are attached. An 8.5×11 inch sheet of notes (both sides) is permitted. The Casio 991 calculator is permitted. HOWEVER, do reasonable simplifications.
- Write the answers in the space provided, continue on the backs of pages if needed.
- SHOW YOUR WORK CLEARLY. Correct answers without clear work showing how you got there will not receive full marks. Marks per part question are shown in brackets at the right margin.

Marks: Please do not write in the space below.

Problem 1 [10]	Problem 4 [10]
Problem 2 [10]	Problem 5 [10]
Problem 3 [10]	Problem 6 [10]

Total: [60]

1. Let U and V be independent, discrete random variables, each uniformly distributed on the integers $1, \ldots, n$, i.e., $P(U = i) = P(V = i) = 1/n$, for $i = 1, \ldots, n$. Let $X = U - V$ and $Y = U + V$.

(a) Find $p_X(x)$ and $p_Y(y)$, the marginal pmfs of X and Y, respectively. *Note:* This is an exercise in constraints! [7]

Solution: The possible values of X are $-(n-1), \ldots, n-1$. For $x \in \{0, \ldots, n-1\}$ there are $n - x$ pairs (u, v) such that $u - v = k$, namely $(x+1, 1), \ldots, (n, n-x)$. Each of these pairs has probability $\frac{1}{n^2}$ and so $P(X = x) = \frac{n-x}{n^2}$. For $x \in \{-(n-1), \ldots, -1\}$ there are $n - |x|$ pairs (u, v) such that $u - v = x$, namely $(1, 1+|x|), \ldots, (n-|x|, n)$. Each of these pairs has probability $\frac{1}{n^2}$ and so $P(X = x) = \frac{n-|x|}{n^2}$. We can combine these two cases to write the marginal pmf of X as

$$f_X(x) = \begin{cases} \frac{n-|x|}{n^2} & \text{for } x = -(n-1), \ldots, n-1 \\ 0 & \text{otherwise.} \end{cases}$$

The possible values of Y are $2, \ldots, 2n$. For $y \in \{2, \ldots, n+1\}$ there are $y - 1$ pairs (u, v) such that $u + v = y$, namely $(1, y-1), \ldots, (y-1, 1)$. Each of these pairs has probability $\frac{1}{n^2}$ and so $P(Y = y) = \frac{y-1}{n^2}$. For $y \in \{n+2, \ldots, 2n\}$ there are $2n - (y-1)$ pairs (u, v) such that $u + v = y$, namely $(y-n, n), \ldots, (n, y-n)$. Each of these pairs has probability $\frac{1}{n^2}$ and so $P(Y = y) = \frac{2n-(y-1)}{n^2}$. To summarize, the marginal pmf of Y is

$$f_Y(y) = \begin{cases} \frac{y-1}{n^2} & \text{for } y = 2, \ldots, n+1 \\ \frac{2n-(y-1)}{n^2} & \text{for } y = n+2, \ldots, 2n \\ 0 & \text{otherwise.} \end{cases}$$

(b) Are X and Y independent? Justify your answer. [3]

Solution: X and Y are not independent since, for example, $P(X = n-1, Y = 2) = 0$ but $P(X = n - 1)$ and $P(Y = 2)$ are both positive.

2. Let $X_1, \ldots, X_n$ be independent and identically distributed Exponential(λ) random variables. Compute $E[X_{(1)}e^{-\lambda X_{(2)}}]$, where $X_{(1)}$ and $X_{(2)}$ are the first and second order statistics of $X_1, \ldots, X_n$. [10]

Solution: The joint pdf of $X_{(1)}$ and $X_{(2)}$ is

$$f_{12}(x_1, x_2) = \begin{cases} n(n-1)\lambda^2 e^{-\lambda x_1} e^{-\lambda x_2}(e^{-\lambda x_2})^{n-2} & \text{for } 0 < x_1 < x_2 < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} E[X_{(1)}e^{-\lambda X_{(2)}}] &= n(n-1)\lambda^2 \int_0^\infty \int_{x_1}^\infty x_1 e^{-\lambda x_2} e^{-\lambda x_1} e^{-\lambda(n-1)x_2} dx_2 dx_1 \\ &= (n-1)\lambda \int_0^\infty x_1 e^{-\lambda x_1} e^{-\lambda n x_1} dx_1 \\ &= \frac{n-1}{n+1} \int_0^\infty x_1 \lambda(n+1) e^{-\lambda(n+1)x_1} dx_1 \\ &= \frac{n-1}{\lambda(n+1)^2}, \end{aligned}$$

since the final integral is the mean of an Exponential($\lambda(n+1)$) distribution.

3. We have 5 boxes, which are initially empty, and 4 red balls and 4 blue balls. For each ball, we pick a box at random (equally likely) and place the ball in the box, independently from ball to ball. Let

$$\begin{aligned} W &= \text{the number of empty boxes} \\ X &= \text{the number of boxes with no red balls} \\ Y &= \text{the number of boxes with no blue balls} \\ Z &= \text{the number of boxes with at least one red and one blue ball} \end{aligned}$$

(a) Find $\mathrm{Cov}(W, X)$. [7]

Solution: Write $W = W_1 + \ldots + W_5$ and $X = X_1 + \ldots + X_5$, where W_i is the indicator that box i is empty and X_i is the indicator that box i has no red balls. Then

$$\begin{aligned} E[W_i] &= P(\text{box } i \text{ is empty}) = \left(\frac{4}{5}\right)^8 \\ E[X_i] &= P(\text{box } i \text{ has no red balls}) = \left(\frac{4}{5}\right)^4 \end{aligned}$$

so that $E[W] = 5\left(\frac{4}{5}\right)^8$ and $E[X] = 5\left(\frac{4}{5}\right)^4$. Next,

$$E[W_i X_i] = P(\text{box } i \text{ is empty}) = \left(\frac{4}{5}\right)^8$$

and, for $i \neq j$,

$$E[W_i X_j] = P(\text{box } i \text{ is empty and box } j \text{ has no red balls}) = \left(\frac{3}{5}\right)^4 \left(\frac{4}{5}\right)^4.$$

Then $E[WX] = 5\left(\frac{4}{5}\right)^8 + 20\left(\frac{3}{5}\right)^4\left(\frac{4}{5}\right)^4$, and

$$\begin{aligned} \mathrm{Cov}(W, X) &= E[WX] - E[W]E[X] \\ &= 5\left(\frac{4}{5}\right)^8 + 20\left(\frac{3}{5}\right)^4\left(\frac{4}{5}\right)^4 - 25\left(\frac{4}{5}\right)^8\left(\frac{4}{5}\right)^4 \\ &= .83886 + 1.06168 - 1.71799 = .1826. \end{aligned}$$

(b) Find $\rho(W, X+Y+Z)$, where $\rho(\cdot,\cdot)$ is the correlation coefficient. *Hint:* This is a one (or two) sentence answer. [3]

Solution: Noting that $X+Y+Z-W=5$ (because $X+Y$ counts the number of empty boxes twice), we get that $W = X+Y+Z-5$. Thus W is a linear function of $X+Y+Z$ with a positive coefficient (equal to one), and so the correlation between W and $X+Y+Z$ is equal to one.

4. Suppose 2 urns (urn 1 and urn 2) contain N balls in total (so if urn 1 contains i balls then urn 2 contains $N-i$ balls). Balls are drawn independently. On each draw one ball is drawn and is then placed in the other urn (the urn the ball was not drawn from).

(a) Here assume $N = 4$ and a ball is selected at random (i.e., suppose the balls are numbered 1 to 4 and for each draw ball number i is selected with probability 1/4, for $i = 1, 2, 3, 4$, and then drawn from whichever urn it happens to be in). Let M_i, $i = 0, 1, 2, 3, 4$, denote the expected number of draws until one of the urns is empty, if initially urn 1 contains i balls and urn 2 contains $4-i$ balls. Find M_i, $i = 1, 2, 3$. Note that $M_0 = M_4 = 0$. [5]

Solution: By conditioning on which urn the first drawn ball was drawn from, and using the boundary conditions $M_0 = M_4 = 0$, we obtain the equations

$$M_1 = \frac{1}{4} + \frac{3}{4}(1 + M_2) \tag{1}$$

$$M_2 = \frac{1}{2}(1 + M_1) + \frac{1}{2}(1 + M_3) \tag{2}$$

$$M_3 = \frac{3}{4}(1 + M_2) + \frac{1}{4}, \tag{3}$$

using the boundary conditions $M_0 = M_4 = 0$. Solving, (3) into (2) gives $M_2 = 1 + \frac{1}{2}M_1 + \frac{1}{2}(1 + \frac{3}{4}M_2)$, or $\frac{5}{8}M_2 = \frac{3}{2} + \frac{1}{2}M_1$, or $M_2 = \frac{12}{5} + \frac{4}{5}M_1$. Plugging this into (1) then gives $M_1 = 1 + \frac{3}{4}(\frac{12}{5} + \frac{4}{5}M_1)$, or $\frac{2}{5}M_1 = \frac{14}{5}$, or $M_1 = 7$. Then $M_2 = \frac{40}{5} = 8$ and from (3) we have $M_3 = 1 + \frac{3}{4}M_2$, or $M_3 = 7$ (we could have deduced that $M_1 = M_3$ by symmetry). Summarizing, $M_1 = 7$, $M_2 = 8$, and $M_3 = 7$.

(b) Here assume N is any fixed positive integer but now suppose for each draw we select an *urn* at random (each with probability 1/2) and then select a ball from the chosen urn. Let p_i, $i = 0, \ldots, N$, denote the probability that urn 1 is empty before urn 2 is empty, if initially urn 1 contains i balls and urn 2 contains $N - i$ balls. Find p_i, $i = 1, \ldots, N-1$. Note that $p_0 = 1$ and $p_N = 0$. [5]

Solution: Conditioning, on which urn the first ball was drawn from and using the boundary conditions $p_0 = 1$ and $p_N = 0$, we obtain the system of equations

$$p_i = \frac{1}{2}p_{i-1} + \frac{1}{2}p_{i+1} \quad \text{for } i = 1, \ldots, N-1 \tag{4}$$

To solve these equations, we can define $d_i = p_i - p_{i-1}$, for $i = 1, \ldots, N$. Then the equations (4) become $d_i = d_{i+1}$, for $i = 1, \ldots, N-1$. Because of the boundary conditions we have $d_1 + \ldots + d_N = -1$, or $d_1 = -\frac{1}{N}$. Going back to the p_i we have $d_1 + \ldots + d_i = p_i - 1$, or $p_i = 1 + id_1 = 1 - \frac{i}{N}$.

5. Let X and $X_1, X_2, \ldots$ be random variables each having a $N(0,1)$ distribution. Suppose (X_n, X) has a bivariate normal distribution for each n and the correlation between X_n and X is $\rho(X_n, X) = \rho_n$, for $n \geq 1$.

(a) Show that X_n converges to X in distribution as $n \to \infty$ (for arbitrary correlations ρ_n). [2]

Solution: If Φ is the standard normal cdf and F_n is the cdf of X_n for $n \geq 1$, then $F_n(x) = \Phi(x)$ for all x (this is given). Therefore, $F_n(x) \to \Phi(x)$ as $n \to \infty$ (trivially). It is also given that $\Phi(x)$ is the cdf of X. Therefore, X_n converges to X in distribution.

(b) If $\rho_n \to 1$ as $n \to \infty$, show that X_n converges to X in probability as $n \to \infty$. [4]

Solution: Let $\epsilon > 0$ be given. Both X_n and X are zero mean so $X_n - X$ also has zero mean. So by Chebyshev'e inequality,

$$P(|X_n - X| > \epsilon) \leq \frac{\text{Var}(X_n - X)}{\epsilon^2}. \tag{5}$$

But

$$\text{Var}(X_n - X) = \text{Var}(X_n) + \text{Var}(X) - 2\text{Cov}(X_n, X) = 1 + 1 - 2\rho_n = 2(1 - \rho_n) \tag{6}$$

So if $\rho_n \to 1$ as $n \to \infty$ then $P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$ for any $\epsilon > 0$. Thus, X_n converges to X in probability.

(c) Show that if $\rho_n = 1 - a^n$ for some constant $a \in (0,1)$, then X_n converges to X with probability 1 as $n \to \infty$. Do you get convergence with probability 1 if $a = 0$? If $a = 1$? Prove your answers. [4]

Solution: Plugging $\rho_n = 1 - a^n$ into (6) we have $\mathrm{Var}(X_n - X) = 2a^n$. Plugging this into Chebyshev's inequality in (5) we have

$$P(|X_n - X| > \epsilon) \leq \frac{2a^n}{\epsilon^2}.$$

If $a \in (0,1)$ the sum on the right converges and so X_n converges to X with probability 1 by the sufficient condition from class. If $a = 0$ the sum is again convergent (it is equal to 0) and so X_n converges to X with probability 1. If $a = 1$ then $\rho_n = 0$ so that X_n is independent of X for any n (since (X_n, X) is bivariate normal). In this case $X_n - X$ has a $N(0,2)$ distribution and so $P(|X_n - X| > \epsilon) = 2(1 - \Phi(\epsilon/\sqrt{2}))$, where Φ is the standard normal cdf. This is some positive constant for every n and so $P(|X_n - X| > \epsilon)$ does not converge to 0 as $n \to \infty$. Therefore, X_n does not converge to X in probability. This then implies that X_n does not converge to X with probability 1.

6. Suppose 80 points are uniformly distributed in the ball in $\mathbb{R}^3$ centred at the origin with radius 1. Let D_i be the Euclidean distance from the origin of the ith point, $i = 1, \ldots, 80$, and let $\overline{D} = \frac{1}{80}\sum_{i=1}^{80} D_i$. Use the central limit theorem to find a value c satisfying $P(|\overline{D} - E[\overline{D}]| \le c) = .95$. Note that the volume of a ball of radius r is $4\pi r^3/3$. [10]

Solution: We first find the distribution of D_i. For $x \in (0,1)$, $D_i \le x$ if and only if the ith point in the ball is in the ball in $\mathbb{R}^3$ centred at the origin of radius x. That is, $P(D_i \le x) = \frac{4\pi x^3/3}{4\pi/3} = x^3$. From this we have that the pdf of D_i is

$$f_D(x) = \begin{cases} 3x^2 & \text{for } x \in [0,1] \\ 0 & \text{otherwise.} \end{cases}$$

The mean, second moment, and variance are then easily computed to be

$$E[D_i] = \frac{3}{4}, \quad E[D_i^2] = \frac{3}{5}, \quad \mathrm{Var}(D_i) = \frac{3}{80}.$$

By the central limit theorem

$$\begin{aligned} P(|\overline{D} - 3/4| \le c) &= P\left(\frac{\sqrt{80}|\overline{D} - 3/4|}{\sqrt{3/80}} \le \frac{c\sqrt{80}}{\sqrt{3/80}}\right) \\ &\approx P\left(|Z| \le \frac{c\sqrt{80}}{\sqrt{3/80}}\right), \end{aligned}$$

where Z has a $N(0,1)$ distribution. In order for this probability to be equal to .95 we need $\frac{c\sqrt{80}}{\sqrt{3/80}}$ equal to 1.96. Solving for c gives

$$c = \frac{1.96\sqrt{3}}{80} \approx .0424.$$

Formula Sheet

Special Distributions

Exponential distribution with parameter $\lambda > 0$:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \qquad E[X] = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

$$E[X] = \mu, \quad \mathrm{Var}(X) = \sigma^2$$

The distribution function, $\Phi(z)$, of a standard normal random variable

Note: $\Phi(-z) = 1 - \Phi(z)$ for any $z \in \mathbb{R}$.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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	0.7704	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
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998