Probability Theory
2015/16 Semester IIb
Instructor: Daniel Valesin
Reexamination
5/7/2016
Duration: 3 hours

Name: ____________________
Student number: __________

This exam contains 9 pages (including this cover page) and 7 problems. Enter all requested information on the top of this page.

Your answers should be written in this booklet. Avoid handing in extra paper.

You are allowed to have two hand-written sheets of paper and a calculator.

You are required to show your work on each problem.

Do not write on the table below.

Problem	Points	Score
1	14	
2	10	
3	14	
4	14	
5	14	
6	14	
7	10	
Total:	90	

Standard Normal cumulative distribution function
The value given in the table is $F_X(x)$ for $X \sim \mathcal{N}(0,1)$.

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
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
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

1. (a) (7 points) A deck consists of 52 distinct cards, among which 4 are aces. The deck is distributed randomly among 4 players, so that each player receives 13 cards. Find the probability that each player receives one ace.

   (b) (7 points) 40% of the boys in a certain village are polite and the remaining 60% are impolite. Polite boys open doors to elders 2/3 of the time, whereas impolite boys only do so half the time. Little Paul is seen opening the door to Mrs Marple, but not opening it to Mr Poirot. What is the probability that he is polite?

Solution.

(a) The total number of ways we can distribute hands to the four players is

$$\binom{52}{13} \cdot \binom{39}{13} \cdot \binom{26}{13} \cdot \binom{13}{13} = \frac{52!}{13!39!} \cdot \frac{39!}{13!26!} \cdot \frac{26!}{13!13!} \cdot 1 = \frac{52!}{(13!)^4}.$$

Let us now count the number of assignments of hands in which each player receives one ace. First, there are 4! ways to distribute the aces to the players (so that each player gets one ace). Next, we distribute the remaining $52 - 4 = 48$ cards equally to the players; the number of ways this can be done is

$$\binom{48}{12} \cdot \binom{36}{12} \cdot \binom{24}{12} \cdot \binom{12}{12} = \frac{48!}{12!36!} \cdot \frac{36!}{12!24!} \cdot \frac{24!}{12!12!} \cdot 1 = \frac{48!}{(12!)^4}.$$

The desired probability is thus

$$\frac{4! \cdot 48!/(12!)^4}{52!/(13!)^4} = \frac{4!48!(13!)^4}{52!(12!)^4} = \frac{4!(13)^4}{52 \cdot 51 \cdot 50 \cdot 49}.$$

(b) Define the events

$$P = \{\text{Paul is polite}\},$$

$$I = \{\text{Paul is impolite}\},$$

$$E = \{\text{Paul opens the door to Mrs Marple and does not open it to Mr Poirot}\}.$$

Then,

$$\mathbb{P}(P) = 0.4, \quad \mathbb{P}(I) = 0.6, \quad \mathbb{P}(E|P) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}, \quad \mathbb{P}(E|I) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Hence, by Bayes' formula,

$$\mathbb{P}(P|E) = \frac{\mathbb{P}(E|P) \cdot \mathbb{P}(P)}{\mathbb{P}(E|P) \cdot \mathbb{P}(P) + \mathbb{P}(E|I) \cdot \mathbb{P}(I)} = \frac{\frac{2}{9} \cdot 0.4}{\frac{2}{9} \cdot 0.4 + \frac{1}{4} \cdot 0.6} = 0.372093.$$

2. (10 points) Let $X \sim$ Geometric(1/2), $Y \sim$ Geometric(1/3) and $Z \sim$ Geometric(1/4) be independent. Find $\mathbb{P}(X = Y = Z)$ and $\mathbb{P}(X < Y < Z)$.

Solution.

$$\begin{aligned}
\mathbb{P}(X = Y = Z) &= \sum_{x \geq 1} \mathbb{P}(X = Y = Z = x) \\
&= \sum_{x \geq 1} \left(\frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{4} \cdot \left(\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \right)^{x-1} \right) \\
&= \frac{1}{24} \cdot \sum_{x \geq 1} \left(\frac{1}{4} \right)^{x-1} = \frac{1}{18}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(X < Y < Z) &= \sum_{x \geq 1} \sum_{y > x} \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y) \cdot \mathbb{P}(Z > y) \\
&= \sum_{x \geq 1} \sum_{y > x} \left(\frac{1}{2} \right)^{x} \cdot \frac{1}{3} \cdot \left(\frac{2}{3} \right)^{y-1} \cdot \left(\frac{3}{4} \right)^{y} \\
&= \frac{1}{3} \cdot \left(\frac{2}{3} \right)^{-1} \cdot \sum_{x \geq 1} \left(\frac{1}{2} \right)^{x} \sum_{y > x} \left(\frac{1}{2} \right)^{y} \\
&= \frac{1}{2} \cdot \sum_{x \geq 1} \left(\frac{1}{2} \right)^{x} \cdot \left(\frac{1}{2} \right)^{x} = \frac{1}{6}.
\end{aligned}$$

3. Suppose that we put together in an urn one ball with the number 1 written on it, two balls with the number 2 written on them, ..., n balls with the number n written on them. (The urn ends up containing $1 + 2 + \cdots + n$ balls). We then select a ball at random from the urn (that is, all the balls have the same probability of being chosen). Let Y_n be the number written on the chosen ball.

   (a) (7 points) Find $\mathbb{E}(Y_n)$. You may use the formulas:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}, \quad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

   (b) (7 points) Show that, as $n \to \infty$, $\frac{Y_n}{n}$ converges in distribution to a continuous random variable Y with probability density function

$$f_Y(y) = \begin{cases} 2y, & \text{if } 0 < y < 1; \\ 0, & \text{otherwise.} \end{cases}$$

Solution. The probability of choosing a ball labeled k is

$$f_{Y_n}(k) = \frac{\#\{\text{balls labeled } k\}}{1 + 2 + \cdots + n} = \frac{2k}{n(n+1)}, \qquad k \in \{1, \ldots, n\}.$$

Hence,

$$\mathbb{E}(Y_n) = \sum_{k=1}^{\infty} k \cdot f_{Y_n}(k) = \sum_{k=1}^{n} \frac{2k^2}{n(n+1)} = \frac{2n+1}{3}.$$

Next, note that

$$M_{Y_n/n}(t) = M_{Y_n}(t/n) = \mathbb{E}(e^{tY_n/n}) = \frac{2n}{n+1} \cdot \frac{1}{n} \cdot \sum_{k=1}^{n} \frac{k}{n} e^{tk/n} \xrightarrow{n \to \infty} 2 \int_0^1 y e^{ty}\, dy = M_Y(t).$$

4. Let $X_1, \dots, X_n$ be independent and identically distributed continuous random variables with finite expectation μ and variance σ^2. Let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

   (a) (7 points) Show that
$$f_{\overline{X}_n}(x) = n f_{X_1+\cdots+X_n}(nx).$$

   (b) (7 points) Find the variance of $5\overline{X}_n - 4X_1$.

Solution.

(a)
$$F_{\overline{X}_n}(x) = \mathbb{P}\left(\frac{1}{n}(X_1+\cdots+X_n) \le x\right) = F_{X_1+\cdots+X_n}(nx),$$

so
$$f_{\overline{X}_n}(x) = \tfrac{d}{dx}F_{\overline{X}_n}(x) = \tfrac{d}{dx}F_{X_1+\cdots+X_n}(nx) = n f_{X_1+\cdots+X_n}(nx).$$

(b)
$$\begin{aligned}
\mathrm{Var}(5\overline{X}_n - 4X_1) &= \mathrm{Var}\left(\frac{5}{n}X_1 + \cdots + \frac{5}{n}X_n - 4X_1\right) \\
&= \mathrm{Var}\left(\left(-4+\frac{5}{n}\right)X_1 + \frac{5}{n}X_2 + \cdots + \frac{5}{n}X_n\right) \\
&= \left(-4+\frac{5}{n}\right)^2 \mathrm{Var}(X_1) + \left(\frac{5}{n}\right)^2 \mathrm{Var}(X_2) + \cdots + \left(\frac{5}{n}\right)^2 \mathrm{Var}(X_n) \\
&= \left(\frac{25}{n^2} + 16 - \frac{40}{n}\right)\sigma^2 + \frac{25(n-1)}{n^2}\sigma^2 = \left(16 - \frac{15}{n}\right)\sigma^2.
\end{aligned}$$

5. Let X and Y be independent random variables, both following the exponential distribution with parameter 1.

   (a) (7 points) Find the joint probability density function of $U = X - Y$ and $V = X + 2Y$.

   (b) (7 points) Find $f_{U|V}$.

Solution.

(a) We have

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y}, & \text{if } x, y \geq 0; \\ 0 & \text{otherwise} \end{cases}$$

and $(U, V) = g(X, Y)$, where $g(x, y) = (x - y, x + 2y)$. Letting

$$D = \{(x, y) : x, y \geq 0\},$$

we have

$$R = g(D) = \{(u, v) : v \geq 0,\ -v/2 \leq u \leq v\}$$

(this can be found by for example drawing the lines $(0, \infty) \ni s \mapsto g(s, y)$ for each $y \geq 0$). Moreover,

$$h(u, v) = g^{-1}(u, v) = \frac{1}{3}(2u + v, v - u) \quad \Longrightarrow \quad J(u, v) = \det \begin{pmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{pmatrix} = \frac{1}{3}.$$

This gives, for $(u, v) \in R$,

$$f_{U,V}(u, v) = f_{X,Y}(h(u, v)) \cdot |J(u, v)| = \frac{1}{3} \cdot e^{-\frac{1}{3}(2u+v)} \cdot e^{-\frac{1}{3}(v-u)} = \frac{1}{3} \cdot e^{-\frac{1}{3}(u+2v)}.$$

In conclusion,

$$f_{U,V}(u, v) = \begin{cases} \frac{1}{3} e^{-\frac{u+2v}{3}}, & \text{if } v \geq 0,\ -v/2 \leq u \leq v, \\ 0 & \text{otherwise.} \end{cases}$$

(b) For $v > 0$ and $u \in (-v/2, v)$,

$$f_{U|V}(u \mid v) = \frac{\frac{1}{3} \cdot e^{-\frac{1}{3}(u+2v)}}{\int_{-v/2}^{v} \frac{1}{3} \cdot e^{-\frac{1}{3}(u+2v)}\, du} = \frac{e^{-\frac{1}{3}(u+2v)}}{3(e^{-v/2} - e^{-v})}.$$

6. (a) (7 points) Let

$$\begin{array}{c} A_1^{(1)},\ A_2^{(1)},\ A_3^{(1)},\ldots, \\ A_1^{(2)},\ A_2^{(2)},\ A_3^{(2)},\ldots, \\ \vdots \\ A_1^{(k)},\ A_2^{(k)},\ A_3^{(k)},\ldots \end{array}$$

be k sequences of events in a probability space. Show that, if

$$\mathbb{P}(A_n^{(i)}) \xrightarrow{n\to\infty} 1 \quad \text{for } i = 1, 2, \ldots, k,$$

then

$$\mathbb{P}\left(\bigcap_{i=1}^{k} A_n^{(i)}\right) \xrightarrow{n\to\infty} 1.$$

(b) (7 points) Let X_1, X_2, X_3, ... be a sequence of independent and identically distributed discrete random variables, all with the same probability mass function f. Show that, for any real numbers $x_1, \ldots, x_k$ and any $\varepsilon > 0$, we have

$$\mathbb{P}\left(\left|\frac{\#\{m \le n : X_m = x_i\}}{n} - f(x_i)\right| < \varepsilon \text{ for all } i \in \{1, \ldots, k\}\right) \xrightarrow{n\to\infty} 1.$$

Hint. Use part (a).

Solution.

(a)

$$1 \ge \mathbb{P}\left(\bigcap_{i=1}^{k} A_n^{(i)}\right) = 1 - \mathbb{P}\left(\bigcup_{i=1}^{k} (A_n^{(i)})^c\right) \ge 1 - \sum_{i=1}^{k} \mathbb{P}((A_n^{(i)})^c) = 1 - \sum_{i=1}^{k}\left(1 - \mathbb{P}(A_n^{(i)})\right) \xrightarrow{n\to\infty} 1.$$

(b) Let $Y_m^{(i)} = \mathbb{1}_{\{X_m = x_i\}}$. Then, for any i, the random variables $Y_1^{(i)}$, $Y_2^{(i)}$, $Y_3^{(i)}$, ... are independent and distributed as

$$f_{Y_m^{(i)}}(1) = 1 - f_{Y_m^{(i)}}(0) = \mathbb{P}(X_m = x_i) = f(x_i),$$

that is, $Y_m^{(i)} \sim \text{Bernoulli}(f(x_i))$. Then, by the Weak Law of Large Numbers,

$$\frac{\sum_{m=1}^{n} Y_m^{(i)}}{n} \xrightarrow[\mathbb{P}]{n\to\infty} \mathbb{E}(Y_1^{(i)}) = f(x_i),$$

that is, for any $\varepsilon > 0$, defining

$$A_n^{(i)} = \left\{\left|\frac{\sum_{m=1}^{n} Y_m^{(i)}}{n} - f(x_i)\right| < \varepsilon\right\} = \left\{\left|\frac{\#\{m \le n : X_m = x_i\}}{n} - f(x_i)\right| < \varepsilon\right\},$$

we have $\mathbb{P}(A_n^{(i)}) \xrightarrow{n\to\infty} 1$ for $i = 1, \ldots, k$. The desired result now follows from part (a).

7. (10 points) An exam is applied to two classes of 100 students each. The scores of students of class 1 are identically distributed with expectation 7.5 and variance 6. The scores of students of class 2 are identically distributed with expectation 8 and variance 3. Assume that students' scores are all independent. Estimate the probability that the average exam score of class 1 is larger than the average exam score of class 2.

   Solution. Letting $X_1, \ldots, X_{100}$ denote the exam scores of class 1 and $Y_1, \ldots, Y_{100}$ the exam scores of class 2, by the Central Limit Theorem we have

$$\bar{X}_{100} = \frac{1}{100} \sum_{i=1}^{100} X_i \approx \mathcal{N}(7.5, 6/100), \qquad \bar{Y}_{100} = \frac{1}{100} \sum_{i=1}^{100} Y_i \approx \mathcal{N}(8, 3/100).$$

   Then, since $\bar{X}_{100}$ and $\bar{Y}_{100}$ are independent,

$$\bar{X}_{100} - \bar{Y}_{100} \approx \mathcal{N}(7.5 - 8, (6+3)/100) = \mathcal{N}(-.5, 0.09).$$

   Hence, letting $Z \sim \mathcal{N}(0, 1)$,

$$\mathbb{P}(\bar{X}_{100} - \bar{Y}_{100} > 0) = \mathbb{P}\left(Z > \frac{0.5}{0.3}\right) \approx 1 - F_Z(1.66) \approx 0.0485.$$