Probability Theory
2016/17 Semester IIb
Instructor: Daniel Valesin
Reexamination
13/7/2017
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	12	
2	14	
3	12	
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) (5 points) We place 9 distinct books from left to right in a shelf. Find the number of possible arrangements in which books A and B are **not** adjacent to each other.

   (b) (7 points) A machine produces 6-sided dice. The machine is defective: while 99.9% of the dice it produces are normal, the remaining 0.1% have all their faces marked 6. Suppose I take (at random) a die produced by this machine and roll it n times, and then I inform you that all the rolls resulted in 6. For which values of n do you now think it is more likely that I took a defective die than that I took a normal die?

Solution. (a) With no restriction, the number of arrangements is 9!. The number of arrangements in which A, B are adjacent is $2 \cdot 8!$ (to see this, pretend A and B are a single book, so there are 8 books on total, but then we multiply by 2 because we can have either A before B or B before A). Hence, the answer is

$$9! - 2 \cdot 8! = 282240.$$

(b)

$$\mathbb{P}(\text{Defective die} \mid n \text{ times } 6) = \frac{0.001 \cdot 1}{0.001 \cdot 1 + 0.999 \cdot (1/6)^n} = \frac{1}{1 + 999/6^n}.$$

This is larger than $\frac{1}{2}$ if $6^n > 999$, which holds for $n \geq 4$.

2. (a) (7 points) Let $X_1, \dots, X_n$ be independent $\mathcal{N}(0,1)$ random variables. Let

$$Y_1 = \left| \frac{1}{n} \sum_{i=1}^{n} X_i \right|, \qquad Y_2 = \frac{1}{n} \sum_{i=1}^{n} |X_i|.$$

Find $\mathbb{E}(Y_1)$ and $\mathbb{E}(Y_2)$.

(b) (7 points) Assume X follows the exponential distribution with parameter 1. Find $\mathbb{E}(\lfloor X \rfloor)$. Here $\lfloor \cdot \rfloor$ denotes the floor function: for $x \in [0, \infty)$, $\lfloor x \rfloor$ is the largest integer smaller than or equal to x.

Solution. (a) $\frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(0, 1/n)$, so

$$\mathbb{E}(Y_1) = \int_{-\infty}^{\infty} |x| \cdot \sqrt{\frac{n}{2\pi}} e^{-\frac{nx^2}{2}} \; d = 2\int_0^{\infty} x \cdot \sqrt{\frac{n}{2\pi}} e^{-\frac{nx^2}{2}} \; dx = \sqrt{\frac{2}{\pi n}},$$

integrating by substitution. Next, note that

$$\mathbb{E}(|X_i|) = \int_{-\infty}^{\infty} |x| \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \; dx = \sqrt{\frac{2}{\pi}} \int_0^{\infty} x \cdot e^{-\frac{x^2}{2}} \; dx = \sqrt{\frac{2}{\pi}}$$

again by substitution, so

$$\mathbb{E}(Y_2) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(|X_i|) = \frac{1}{n} \cdot n \cdot \sqrt{\frac{2}{\pi}} = \sqrt{\frac{2}{\pi}}.$$

(b)

$$\mathbb{E}\left(\lfloor X \rfloor\right) = \int_{-\infty}^{\infty} \lfloor x \rfloor \cdot f_X(x) \; dx = \sum_{k=0}^{\infty} k \cdot (e^{-k} - e^{-k-1}).$$

For each $N \in \mathbb{N}$, we have

$$\begin{aligned} \sum_{k=0}^{N} k \cdot (e^{-k} - e^{-k-1}) &= 0 \cdot (1 - e^{-1}) + 1 \cdot (e^{-1} - e^{-2}) + 2 \cdot (e^{-2} - e^{-e}) + \cdots + N \cdot (e^{-N} - e^{-N-1}) \\ &= e^{-1} + e^{-2} + \cdots + e^{-N} - e^{-N-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{k=0}^{\infty} k \cdot (e^{-k} - e^{-k-1}) &= \lim_{N\to\infty} \sum_{k=0}^{N} k \cdot (e^{-k} - e^{-k-1}) \\ &= \lim_{N\to\infty} e^{-1} + e^{-2} + \cdots + e^{-N} - e^{-N-1} \\ &= \sum_{k=1}^{\infty} (1/e)^k = \frac{1/e}{1 - (1/e)} = \frac{1}{e-1}. \end{aligned}$$

3. In a field, hunters attempt to shoot ducks which are flying by. Assume that:

- there are h hunters and d ducks;
- each hunter chooses a duck uniformly at random and shoots, hitting (and killing) it with probability $p \in (0,1)$;
- the choices and shots of all hunters are independent.

Let D be the number of ducks which die.

(a) (5 points) Find $\mathbb{E}(D)$.

(b) (4 points) Let A be the event that each duck is chosen by at most one hunter. Show that, when h and p are kept fixed and d tends to infinity, $\mathbb{P}(A)$ tends to 1.

(c) (3 points) Prove that, when h and p are kept fixed and d tends to infinity, D converges in distribution to a random variable X; identify the distribution of X.

Hint. For $m \in \{0,\dots,h\}$, write:
$\mathbb{P}(D=m) = \mathbb{P}(\{D=m\} \mid A)\cdot\mathbb{P}(A) + \mathbb{P}(\{D=m\} \mid A^c)\cdot\mathbb{P}(A^c)$.

Solution. (a) Let $X_i = \mathbb{1}\{\text{Duck } i \text{ survives}\}$. We have

$$\mathbb{E}(X_i) = \mathbb{P}(\text{Duck } i \text{ survives}) = \prod_{j=1}^{h} \mathbb{P}\left(\begin{array}{l}\text{hunter } j \text{ does not}\\ \text{kill duck } i\end{array}\right) = \left(\frac{d-1}{d} + \frac{1}{d}(1-p)\right)^h = \left(1-\frac{p}{d}\right)^h.$$

$$\implies \mathbb{E}(D) = \mathbb{E}\left(d - \sum_{i=1}^{d} X_i\right) = d - d\left(1-\frac{p}{d}\right)^h = d\left(1-\left(1-\frac{p}{d}\right)^h\right).$$

(b)

$$1 \le k,\ell \le h,\ k \neq \ell \implies \mathbb{P}\left(\begin{array}{c}\text{hunters } k \text{ and } \ell\\ \text{choose same duck}\end{array}\right) = \sum_{i=1}^{d} \mathbb{P}\left(\begin{array}{c}\text{hunters } k \text{ and } \ell\\ \text{choose duck } i\end{array}\right) = d\cdot\left(\frac{1}{d}\right)^2 = \frac{1}{d}.$$

$$\implies \mathbb{P}(A^c) = \mathbb{P}\left(\bigcup_{k\in\{1,\dots,h\}} \bigcup_{\substack{\ell\in\{1,\dots,h\}\\ \ell\neq k}} \left\{\begin{array}{c}\text{hunters } k \text{ and } \ell\\ \text{choose same duck}\end{array}\right\}\right) \le \frac{h^2}{d} \xrightarrow{d\to\infty} 0,$$

so $\mathbb{P}(A) = \lim_{d\to\infty}(1-\mathbb{P}(A^c)) = 1$.

(c) Each hunter gets an independent trial to kill its chosen duck, and in case A occurs, the chosen ducks are all distinct, so, if $d \ge h$,

$$\mathbb{P}(\{D=m\} \mid A) = \binom{h}{m} p^m (1-p)^{h-m}, \qquad m \in \{0,\dots,h\}.$$

Hence,

$$f_D(m) = \mathbb{P}(\{D=m\}) = \mathbb{P}(\{D=m\} \mid A)\cdot\mathbb{P}(A) + \mathbb{P}(\{D=m\} \mid A^c)\cdot\mathbb{P}(A^c),$$

so we see that D converges in distribution to $X \sim \text{Bin}(h,p)$:

$$\left| f_D(m) - \binom{h}{m} p^m (1-p)^{h-m} \right| \le |\mathbb{P}(A) - 1| \cdot \binom{h}{m} p^m (1-p)^{h-m} + \mathbb{P}(A^c) \xrightarrow{d\to\infty} 0.$$

4. Let X be a random variable following the Poisson(λ) distribution. Let Y be a random variable such that, for each $n \in \{0, 1, 2, \ldots\}$, the probability that Y equals n is proportional to $n^2 \cdot f_X(n)$.
   (a) (7 points) Find the probability mass function of Y.
   (b) (7 points) Find the moment-generating function of Y.

Solution. (a) For $n \geq 0$, we have $f_Y(n) = C \cdot n^2 \cdot \frac{\lambda^n}{n!} e^{-\lambda}$, so

$$C = \left(\sum_{n=1}^{\infty} n^2 \cdot \frac{\lambda^n}{n!} e^{-\lambda} \right)^{-1} = \left(\mathbb{E}(X^2) \right)^{-1} = \left(\mathrm{Var}(X) + \mathbb{E}(X)^2 \right)^{-1} = \frac{1}{\lambda + \lambda^2}.$$

We thus have

$$f_Y(n) = \frac{1}{\lambda + \lambda^2} \cdot \frac{\lambda^n \cdot n^2}{n!} \cdot e^{-\lambda}, \qquad n = 1, 2, \ldots.$$

(b)

$$M_Y(t) = \sum_{n=1}^{\infty} e^{tn} \cdot \frac{1}{\lambda + \lambda^2} \cdot \frac{\lambda^n \cdot n^2}{n!} \cdot e^{-\lambda} = \frac{e^{-\lambda}}{\lambda + \lambda^2} \sum_{n=1}^{\infty} \frac{n^2 (\lambda e^t)^n}{n!}$$
$$= \frac{e^{-\lambda}}{\lambda + \lambda^2} \cdot e^{\lambda e^t} \cdot \sum_{n=1}^{\infty} \frac{n^2 (\lambda e^t)^n}{n!} \cdot e^{-\lambda e^t}.$$

Observe that the sum on the right-hand side equals $\mathbb{E}(W^2)$ for $W \sim \mathrm{Poisson}(\lambda e^t)$, so that

$$\sum_{n=1}^{\infty} \frac{n^2 (\lambda e^t)^n}{n!} \cdot e^{-\lambda e^t} = \mathbb{E}(W^2) = \mathrm{Var}(W) + \mathbb{E}(W)^2 = \lambda e^t + (\lambda e^t)^2 = \lambda e^t (\lambda e^t + 1).$$

Putting things together,

$$M_Y(t) = \frac{e^{t - \lambda + \lambda e^t} (\lambda e^t + 1)}{\lambda + 1}.$$

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

   (a) (7 points) Let $Z = X + 2Y$. Find $f_{Z,Y}$, the joint probability density function of Z and Y.

   (b) (7 points) Let U be the smallest of the two values X, Y (that is, $U = \min(X, Y)$) and V be the largest of the two values X, Y (that is, $V = \max(X, Y)$). Find $F_{U,V}$, the joint cumulative distribution function of U and V.

Solution. (a) $g(x, y) = (z, y) = (x + 2y, y)$, so $h = g^{-1}(z, y) = (z - 2y, y)$. The vector (X, Y) takes values on $D = (0, \infty)^2$, and it is easy to see that $R = g(D) = \{(z, y) : z > 0,\ 0 < y < z/2\}$. Then,

$$f_{Z,Y}(z, y) = f_{X,Y}(h(z, y)) \cdot \det \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = e^{-(z-2y)} \cdot e^{-y} = e^{-z+y}, \qquad z > 0,\ 0 < y < z/2.$$

(b) If $u \leq 0$ or $v \leq 0$, we have $F_{U,V}(u, v) = 0$. Assume $u, v > 0$. If $u \geq v$, we have

$$F_{U,V}(u, v) = \mathbb{P}(\min(X, Y) \leq u,\ \max(X, Y) \leq v) = \mathbb{P}(X, Y \leq v) = \mathbb{P}(X \leq v) \cdot \mathbb{P}(Y \leq v) = e^{-2v}.$$

If $0 < u < v$, we have

$$\begin{aligned} F_{U,V}(u, v) &= \mathbb{P}(\min(X, Y) \leq u,\ \max(X, Y) \leq v) \\ &= \mathbb{P}(X \leq u,\ Y \leq v) + \mathbb{P}(u < X < v,\ Y \leq u) \\ &= (1 - e^{-u}) \cdot (1 - e^{-v}) + (e^{-u} - e^{-v}) \cdot (1 - e^{-u}) \\ &= 1 + 2e^{-u-v} - e^{-2u} - 2e^{-v}. \end{aligned}$$

6. (a) (7 points) Let $X_1, \dots, X_n$ be independent Bernoulli(p) random variables. Prove, using moment-generating functions, that $\sum_{i=1}^n X_i$ follows the Binomial(n, p) distribution.

   (b) (7 points) For $p \in (0, 1)$, $\varepsilon > 0$ and $n \in \mathbb{N}$, let

$$I(p, \varepsilon, n) = \{i \in \mathbb{N} : (1 - \varepsilon)np \leq i \leq (1 + \varepsilon)np\}.$$

   Prove that, for fixed p and ε,

$$\lim_{n\to\infty} \sum_{i \in I(p,\varepsilon,n)} \binom{n}{i} p^i (1 - p)^{n-i} = 1.$$

Solution. (a)

$$M_{X_i}(t) = \mathbb{E}(e^{tX_i}) = pe^t + (1 - p)e^0 = 1 + p(e^t - 1),$$

so, using independence,

$$M_{X_1+\cdots M_n} = \prod_{i=1}^n \mathbb{E}(e^{tX_i}) = (1 + p(e^t - 1))^n,$$

which is the moment-generating function of a Binomial(n, p) random variable.

(b)

$$\sum_{i \in I(p,\varepsilon,n)} \binom{n}{i} p^i (1 - p)^{n-i} = \mathbb{P}(|Z_n - np| \leq \varepsilon np),$$

where $Z_n \sim \text{Bin}(n, p)$. By part (a), we can replace Z_n by $\sum_{i=1}^n X_i$, so the above becomes

$$\mathbb{P}\left(\left|\sum_{i=1}^n X_i - np\right| \leq \varepsilon np\right) = \mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i}{n} - p\right| \leq \varepsilon p\right) \xrightarrow{n\to\infty} 1$$

by the Weak Law of Large Numbers.

7. (10 points) In an exam, students' scores are between 0 and 100. Assume that the score that any individual student obtains in the exam is a random variable with mean 70 and standard deviation 9, and that scores of distinct students are independent. An instructor gives the exam to two classes, one of 55 students and another of 90 students. Estimate the probability that the average score in the class with 90 students exceeds that of the other class by at least 2 points. If your calculator doesn't find square roots, you may use the approximations: $\sqrt{2} \approx 1.41$, $\sqrt{5} \approx 2.24$, $\sqrt{11} \approx 3.32$, $\sqrt{29} \approx 5.39$.

Solution. Let $X_1, \ldots, X_{55}$ be the scores in the first class and $Y_1, \ldots, Y_{90}$ be the scores in the second class. The distribution of $\bar{X}_{55}$ is close to that of $U \sim \mathcal{N}(70, 81/55)$, and the distribution of $\bar{Y}_{90}$ is close to that of $V \sim \mathcal{N}(70, 81/90)$. Hence,

$$\mathbb{P}(\bar{Y}_{90} > \bar{X}_{55} + 2) \approx \mathbb{P}(V > U + 2) = \mathbb{P}(V - U > 2).$$

Since U and V can be assumed independent,

$$V - U \sim \mathcal{N}\left(0, \frac{81}{90} + \frac{81}{55}\right).$$

We then have

$$\mathbb{P}(V - U > 2) = \mathbb{P}\left(\frac{V-U}{\sqrt{\frac{81}{90} + \frac{81}{55}}} > \frac{2}{\sqrt{\frac{81}{90} + \frac{81}{55}}}\right) = \mathbb{P}\left(Z > \frac{2}{\sqrt{\frac{81}{90} + \frac{81}{55}}}\right),$$

where $Z \sim \mathcal{N}(0,1)$. Using a calculator, $\frac{2}{\sqrt{\frac{81}{90}+\frac{81}{55}}} \approx 1.298$; the table for the cdf of $\mathcal{N}(0,1)$ then gives

$$\mathbb{P}\left(Z > 1.298\right) = 1 - \mathbb{P}(Z \leq 1.298) \approx 9.85\%.$$