Probability Theory
2016/17 Semester IIb
Instructor: Daniel Valesin
Final Exam
22/6/2017
Duration: 3 hours

Name: ______________________________
Student number: ____________

This exam contains 10 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	14	
3	14	
4	14	
5	10	
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) In how many ways can we distribute m indistinguishable candies to k distinguishable children, assuming $m \geq k$ and each child gets at least one candy?

   (b) (7 points) Half the lamps produced by a factory are of good quality and half of bad quality. The lifetime (in months) of a good-quality lamp is an exponential random variable with parameter $\beta = 2$, and the lifetime of a bad-quality lamp is an exponential random variable with $\beta = 1$. Assume that you pick a lamp at random in the factory, install it and, after t months, it is still working. Let $p(t)$ be the probability that you now attribute to having picked a good-quality lamp. Find the limit of $p(t)$ as $t \to \infty$.

Solution.

(a) A valid assignment of candies to children is uniquely described by a permutation of symbols: '$\circ$' ($m - k$ times) and '$|$' ($k - 1$ times). The number of such permutations is

$$\frac{(m-k+k-1)!}{(m-k)!(k-1)!} = \frac{(m-1)!}{(m-k)!(k-1)!}.$$

(b)

$$A = \{\text{Picked good lamp}\}, \qquad B_t = \{\text{Lamp survived for time} > t\}.$$

$$\begin{aligned} p(t) = \mathbb{P}(A|B_t) &= \frac{\mathbb{P}(B_t|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B_t|A) \cdot \mathbb{P}(A) + \mathbb{P}(B_t|A^c) \cdot \mathbb{P}(A^c)} \\ &= \frac{e^{-t/2} \cdot \frac{1}{2}}{e^{-t/2} \cdot \frac{1}{2} + e^{-t} \cdot \frac{1}{2}} = \frac{1}{1 + e^{-t/2}} \xrightarrow{t \to \infty} 1. \end{aligned}$$

2. (a) (7 points) The discrete random variables X and Y are jointly distributed as follows. X follows the Poisson(λ) distribution, and

$$f_{Y|X}(y|x) = \binom{x}{y} p^y (1-p)^{x-y}, \qquad x \in \{0, 1, \ldots, \},\ y \in \{0, \ldots, x\},$$

where $p \in (0, 1)$. Show that the distribution of Y is Poisson(λp).

(b) (7 points) The continuous random variables Z and W are independent, with Z following the exponential distribution with parameter 1 and W following the (continuous) uniform distribution on $(0, 1)$. Find $\mathbb{P}(Z < W < 3Z)$.

Solution.

(a) For $y \in \{0, 1, 2, \ldots\}$,

$$f_Y(y) = \sum_{x=y}^{\infty} f_X(x) \cdot f_{Y|X}(y|x) = \sum_{x=y}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} \cdot \frac{x!}{y!(y-x)!} \cdot p^y (1-p)^{x-y}$$

$$= \frac{\lambda^y}{y!} \cdot e^{-\lambda} \cdot p^y \cdot \underbrace{\sum_{x=y}^{\infty} \frac{\lambda^{x-y}}{(x-y)!} \cdot (1-p)^{x-y}}_{e^{\lambda(1-p)}} = \frac{(\lambda p)^y}{y!} \cdot e^{-\lambda p}.$$

(b)

$$\mathbb{P}(Z < W < 3Z) = \int_0^{\frac{1}{3}} \int_z^{3z} e^{-z}\, dwdz + \int_{\frac{1}{3}}^{1} \int_z^{1} e^{-z}\, dwdz$$

$$= \int_0^{\frac{1}{3}} 2ze^{-z} dz + \int_{\frac{1}{3}}^{1} (1-z)e^{-z}\, dz$$

$$= -3e^{-\frac{1}{3}} + 2 + e^{-1}.$$

3. Let A be a set with n elements. Assume that we choose a subset of A at random according to the rule that the probability that a certain subset A' is chosen is proportional to the number of elements of A'. Let X be the number of elements of the subset we choose.

   (a) (7 points) Find the probability mass function of X.

   (b) (7 points) Prove that

$$M_X(t) = e^t \left(\frac{e^t+1}{2}\right)^{n-1}, \qquad t \in \mathbb{R}.$$

Solution.

(a) For $A' \subset A$,

$$\mathbb{P}(A' \text{ is chosen}) = C \cdot \#A'.$$

To find C, we compute

$$1 = \sum_{A' \subset A} C \cdot \#A' = C \sum_{k=0}^{n} \sum_{A' : \#A' = k} \#A' = C \sum_{k=0}^{n} \binom{n}{k} \cdot k = 2^n C \underbrace{\sum_{k=0}^{n} \binom{n}{k} \cdot k \cdot \frac{1}{2^n}}_{\mathbb{E}(Y) \text{ for } Y \sim \text{Bin}(n,1/2)} = n2^{n-1}C,$$

so $C = 1/(n2^{n-1})$. Hence,

$$f_X(k) = \sum_{A' : \#A' = k} \mathbb{P}(A' \text{ is chosen}) = \frac{1}{n2^{n-1}} \sum_{A' : \#A' = k} k = \frac{1}{n2^{n-1}} \cdot \binom{n}{k} \cdot k, \qquad k \in \{1, \ldots, n\}.$$

(b)

$$\begin{aligned}
M_X(t) = \sum_{k=1}^{n} e^{tk} \cdot f_X(k) &= \frac{1}{n2^{n-1}} \sum_{k=1}^{n} e^{tk} \cdot \binom{n}{k} \cdot k \\
&= \frac{1}{n2^{n-1}} \sum_{k=1}^{n} e^{tk} \cdot \frac{n!}{(k-1)!(n-k)!} \\
&= \frac{1}{n2^{n-1}} \cdot ne^t \cdot \sum_{k=1}^{n} e^{t(k-1)} \cdot \frac{(n-1)!}{(k-1)!(n-k)!} \\
&= \frac{e^t}{2^{n-1}} \cdot \underbrace{\sum_{\ell=0}^{n-1} e^{t\ell} \cdot \frac{(n-1)!}{\ell!(n-1-\ell)!}}_{(e^t+1)^{n-1}} = e^t \left(\frac{e^t+1}{2}\right)^{n-1}.
\end{aligned}$$

4. (a) (7 points) A device has lifetime denoted by T, which follows an exponential distribution with parameter $\beta = 1.5$. The device has value $V = 5$ if it fails before $t = 3$; otherwise, it has value $V = 2T$. Find the cumulative distribution function of V.

   (b) (7 points) Let $X_1, X_2, \dots, X_n$ be independent random variables, all following an exponential distribution with parameter β. Let

$$Y = \max\{X_1, \dots, X_n\},$$

   that is, Y is equal to the largest among the values $X_1, \dots, X_n$. Find the cumulative distribution function of Y.

(a) The values that V can attain are 5 and any number in the interval $[6, \infty)$. We have

$$\mathbb{P}(V = 5) = \mathbb{P}(T \leq 3) = \int_0^3 \frac{1}{1.5} e^{-x/1.5}\, dx = 1 - e^{-2}.$$

If $x \in [6, \infty)$, we have

$$\mathbb{P}(V \leq x) = \mathbb{P}(T \leq x/2) = \int_0^{x/2} \frac{1}{1.5} e^{-x/1.5}\, dx = 1 - e^{-x/3}.$$

Putting these facts together, we obtain

$$F_V(x) = \begin{cases} 0 & \text{if } x < 5; \\ 1 - e^{-2} & \text{if } 5 \leq x < 6; \\ 1 - e^{-x/3} & \text{if } x \geq 6. \end{cases}$$

(b)

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(\max\{X_1, \dots, X_n\} \leq y\right) = \mathbb{P}\left(\cap\{X_i \leq y\}\right) = \left(F_{X_1}(y)\right)^n = \left(1 - e^{-y/\beta}\right)^n.$$

5. Let X be a random variable with expectation μ_X and variance σ_X^2, and Y a random variable with expectation μ_Y and variance σ_Y^2. Let $\rho_{X,Y}$ be the correlation between X and Y. Express the following quantities in terms of μ_X, μ_Y, σ_X, σ_Y, and $\rho_{X,Y}$.

   (a) (5 points) $\mathrm{Var}(X - 3Y)$;

   (b) (5 points) $\mathbb{E}((3X - 5)(2Y + 1))$.

   Solution.

   (a)
   $$\mathrm{Var}(X - 3Y) = \mathrm{Cov}(X - 3Y, X - 3Y) = \sigma_X^2 + 9\sigma_Y^2 - 6\rho_{X,Y}\sigma_X\sigma_Y.$$

   (b) Note that
   $$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) - \mathbb{E}(Y) \quad \Longrightarrow \quad \mathbb{E}(XY) = \rho_{X,Y}\sigma_X\sigma_Y - \mu_X\mu_Y.$$

   Hence,
   $$\mathbb{E}((3X-5)(2Y+1)) = 6\mathbb{E}(XY)-5+3\mathbb{E}(X)-10\mathbb{E}(Y) = 6(\rho_{X,Y}\sigma_X\sigma_Y-\mu_X\mu_Y)-5+3\mu_X-10\mu_Y.$$

6. (a) (7 points) Let $X_1, X_2, \dots$ be independent and identically distributed random variables with finite expectation and variance. Let $Z_i = X_i + X_{i+1}$, for $i = 1, 2, \dots$. Show that $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^n Z_i$ converges in probability to a constant, and identify this constant.

(b) (7 points) Let $Y_1, Y_2, \dots$ be independent and identically distributed *positive* random variables with finite expectation μ and finite variance σ^2. For each natural number n, let W_n be the largest value of k for which the following inequality holds:

$$\sum_{i=1}^k Y_i \le n.$$

Show that W_n/n converges in probability to $1/\mu$.

Hint. Use the facts that $\{W_n > x\} \subseteq \{\sum_{i=1}^{\lfloor x \rfloor} Y_i < n\}$, $\{W_n < x\} \subseteq \{\sum_{i=1}^{\lceil x \rceil} Y_i > n\}$.

Solution.

(a) For each i, $\mathbb{E}(Z_i) = \mathbb{E}(X_i) + \mathbb{E}(X_{i+1}) = 2\mathbb{E}(X_i) =: \mu_Z$. Let us show that $\bar{Z}_n$ converges in probability to μ_Z as $n \to \infty$. We start computing

$$\mathrm{Var}(\bar{Z}_n) = \frac{1}{n^2} \cdot \mathrm{Var}\left(\sum_{i=1}^n Z_i\right) = \frac{1}{n^2} \cdot \left(\sum_{i=1}^n \mathrm{Var}(X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(Z_i, Z_j)\right). \qquad (\star)$$

Note that $\mathrm{Var}(Z_i) = \mathrm{Var}(X_i) + \mathrm{Var}(X_{i+1}) = 2\mathrm{Var}(X_1)$. Moreover, if $j > i + 1$, then $\mathrm{Cov}(Z_i, Z_j) = 0$, and

$$\mathrm{Cov}(Z_i, Z_{i+1}) = \mathrm{Cov}(X_i + X_{i+1}, X_{i+1} + X_{i+2}) = \mathrm{Var}(X_{i+1}) = \mathrm{Var}(X_1).$$

Hence, the right-hand side of $(\star)$ is equal to

$$\frac{1}{n^2}\left(n \cdot \mathrm{Var}(Z_1) + 2 \cdot (n-1) \cdot \mathrm{Var}(Z_1)\right) = \frac{3n-2}{n^2}\mathrm{Var}(Z_1).$$

Now, for any $\varepsilon > 0$,

$$\mathbb{P}(|\bar{Z}_n - \mu_Z| > \varepsilon) \le \frac{\mathrm{Var}(\bar{Z}_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \cdot \frac{3n-2}{n^2} \xrightarrow{n\to\infty} 0.$$

(b) Fix $\varepsilon > 0$. We have

$$\mathbb{P}\left(\frac{W_n}{n} > \frac{1}{\mu} + \varepsilon\right) = \mathbb{P}\left(W_n > n\left(\frac{1}{\mu} + \varepsilon\right)\right) \le \mathbb{P}\left(\sum_{i=1}^{\left\lfloor n\left(\frac{1}{\mu}+\varepsilon\right)\right\rfloor} Y_i < n\right) = \mathbb{P}\left(\frac{\sum_{i=1}^{a_n} Y_i}{a_n} < \frac{n}{a_n}\right),$$

where $a_n = \left\lfloor n\left(\frac{1}{\mu} + \varepsilon\right)\right\rfloor$. We can find $\varepsilon' > 0$ such that, for n large enough, $\frac{n}{a_n} < \mu - \varepsilon'$, so that

$$\mathbb{P}\left(\frac{\sum_{i=1}^{a_n} Y_i}{a_n} < \frac{n}{a_n}\right) \le \mathbb{P}\left(\frac{\sum_{i=1}^{a_n} Y_i}{a_n} < \mu - \varepsilon'\right) \xrightarrow{n\to\infty} 0$$

by the Weak Law of Large Numbers.

We now turn to

$$\mathbb{P}\left(\frac{W_n}{n} < \frac{1}{\mu} - \varepsilon\right) = \mathbb{P}\left(W_n < n\left(\frac{1}{\mu} - \varepsilon\right)\right) \leq \mathbb{P}\left(\sum_{i=1}^{\left\lceil n\left(\frac{1}{\mu}-\varepsilon\right)\right\rceil} Y_i > n\right) = \mathbb{P}\left(\frac{\sum_{i=1}^{b_n} Y_i}{b_n} > \frac{n}{b_n}\right),$$

where $b_n = \left\lceil n\left(\frac{1}{\mu} - \varepsilon\right)\right\rceil$. We can find $\varepsilon'' > 0$ such that, for n large enough, $\frac{n}{b_n} > \mu + \varepsilon''$, so that

$$\mathbb{P}\left(\frac{\sum_{i=1}^{b_n} Y_i}{b_n} > \frac{n}{b_n}\right) \leq \mathbb{P}\left(\frac{\sum_{i=1}^{b_n} Y_i}{b_n} > \mu + \varepsilon''\right) \xrightarrow{n\to\infty} 0,$$

again by the Weak Law of Large Numbers.

7. (10 points) At each time instant $t = 0, 1, 2, \ldots$, a particle occupies a position in the set $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Assume that the initial position is 0 and, from time t to time $t+1$, the particle moves:

   - one unit to the left with probability $1/3$;
   - one unit to the right with probability $1/6$;
   - two units to the right with probability $1/2$.

   Find a value k such that the probability that the particle is to the left of k at time 10000 is approximately 70%.
   If you don't have a calculator, you may use the approximation: $\sqrt{65} \approx 8.06$.

   Solution. Let $Y_1, Y_2, \ldots$ be independent and identically distributed random variables with probability mass function

   $$f(-1) = \tfrac{1}{3}, \qquad f(1) = \tfrac{1}{6}, \qquad f(2) = \tfrac{1}{2}.$$

   Then, let $W_0 = 0$ and $W_n = \sum_{i=1}^{n} Y_i$ for $n \geq 1$. The sequence $W_0, W_1, \ldots$ thus represents the successive positions of the particle. Note that the expectation and variance of the Y_i's is respectively

   $$\mu = \frac{1}{3}(-1) + \frac{1}{6} \cdot 1 + \frac{1}{2} \cdot 2 = \frac{5}{6},$$
   $$\sigma^2 = \mathbb{E}(Y_i^2) - (\frac{5}{6})^2 = \frac{1}{3} \cdot (-1)^2 + \frac{1}{6} \cdot 1^2 + \frac{1}{2} \cdot 2^2 - \frac{25}{36} = \frac{65}{36}.$$

   We want k such that

   $$\mathbb{P}(W_{10^4} \leq k) \approx 0.7.$$

   By the Central Limit Theorem,

   $$\frac{W_{10^4} - 10^4 \cdot \mu}{10^2 \cdot \sigma} \approx Z \sim \mathcal{N}(0, 1),$$

   so

   $$\mathbb{P}(W_{10^4} \leq k) = \mathbb{P}\left(\frac{W_{10^4} - 10^4 \cdot \mu}{10^2 \cdot \sigma} \leq \frac{k - 10^4 \cdot \mu}{10^2 \cdot \sigma}\right) \approx \mathbb{P}\left(Z \leq \frac{k - 10^4 \cdot \mu}{10^2 \cdot \sigma}\right).$$

   The table for the cumulative distribution function of $\mathcal{N}(0, 1)$ gives

   $$\mathbb{P}(Z \leq 0.53) \approx 0.7,$$

   so we set

   $$\frac{k - 10^4 \cdot \mu}{10^2 \cdot \sigma} = 0.53 \quad \Longrightarrow \quad k = 10^2 \cdot \sqrt{\frac{65}{36}} \cdot 0.53 + 10^4 \cdot \frac{5}{6} \approx 8404.55.$$