This is a take home final exam. You can use your notes, my online notes on canvas and the text book. You are supposed to work on your own text without external help. I'll be available to answer question in person or via email. Please, write clearly and legibly and take a readable scan before uploading.

To solve the Exam problems, I have not collaborated with anyone nor sought external help and the material presented is the result of my own work.

Signature: Sean Eva

Name (print): Sean Eva

Question:	1	2	3	4	5	6	7	Total
Points:	30	15	25	20	10	0	0	100
Score:								

Question:	1	2	3	4	5	6	7	Total
Bonus Points:	0	0	0	0	0	15	15	30
Score:								

Question 1 . . . . *30 point*

Let X be a continuous r.v. with p.d.f. given by

$$f(x) = \begin{cases} p + 2(1-p)x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where p is a parameter.

(a) (10 points) for which value of p is f a valid p.d.f.? (**Hint**: remember that there are 2 conditions you should check.)

In order for this to be a valid p.d.f., it must meet two conditions.
$(1) : \int_{-\infty}^{\infty} f(x)dx = 1$.
Then, $\int_0^1 f(x)dx = 1 : \forall p$. This then means that p can take on any real value.
$(2) : f(x) \geq 0$, for all x, and since X is a continuous random variable, it needs to be piecewise continuous. Therefore, $lim_{0^+} = lim_{0^-1} = 0$, and in order for this to be true $p = 0$.
Therefore, in order for this p.d.f. to be valid for random variable X, p=0. This then redefines $f(x)$ as

$$f(x) \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) (10 points) Compute the expected value $\mathbb{E}(X)$ and the variance $\mathrm{var}(X)$ of X.

$\mathbb{E}(X) = \int_0^1 x * f(x)dx = \int_0^1 2x^2dx = \frac{2}{3}[x^3]_0^1 = \frac{2}{3}$
$var(X) = \int_0^1 x^2 * f(x)dx - (\int_0^1 xf(x)dx)^2 = \int_0^1 2x^3dx - (\int_0^1 2x^2)^2 = \frac{1}{2}[x^4]_0^1 - (\frac{2}{3}[x^3]_0^1)^2 = \frac{1}{2} - (\frac{2}{3})^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$.

(c) (10 points) Show that

$$\mathbb{E}\left(\left(X - \frac{2}{3}\right)^2\right) \geq \left(\frac{p}{6}\right)^2$$

(**Hint**: use Jensen inequality.)

Consider the function $g(X) = (X - \frac{2}{3})^2$. By using the Jensen inequality, we know that $\mathbb{E}(g(X)) \geq g(\mathbb{E}(X))$. Therefore, $\mathbb{E}((X - \frac{2}{3})^2) \geq (\mathbb{E}(X) - \frac{2}{3})^2$. Given that $\mathbb{E}(X) = \frac{2}{3}$, $\mathbb{E}((X - \frac{2}{3})^2) \geq (\frac{2}{3} - \frac{2}{3})^2 = 0$. We also know that $p = 0$, so $(\frac{p}{6})^2 = (\frac{0}{6})^2 = 0$. Therefore, the inequality of $\mathbb{E}((X - \frac{2}{3})^2) \geq (\frac{p}{6})^2$ holds.

Question 2 .... *15 point*

Let N_k, $k = 1, 2, 3, \ldots$, be an infinite sequence of geometric random variable with parameter $p_k = \frac{\lambda}{k}$, that is

$$\mathbb{P}(N_k = n) = (1 - p_k)^{n-1} p_k \qquad \text{for } n \geq 1\,.$$

and $\mathbb{P}(N_k = n) = 0$ for $n < 1$. Moreover let Y be an exponential r.v. with parameter λ, that is

$$f_Y(y) = \lambda e^{-\lambda y} \qquad \text{for } y \geq 0$$

and $f_Y(y) = 0$ for $y < 0$.

Show that $Z_k = N_k/k$ converge in distribution to Y as $k \to \infty$. (**Hint**: compute the c.d.f. of Z_k, that is $F_k(x) = \mathbb{P}(Z_k \leq x)$ for every real number x.)

$$\begin{aligned}
F_k(x) &= \mathbb{P}(Z_k \leq x) \\
&= \mathbb{P}(\frac{N_k}{k} \leq x) \\
&= \mathbb{P}(N_k \leq kx) \\
&= \mathbb{P}(N_k = 1) + \mathbb{P}(N_k = 2) + ... + \mathbb{P}(N_k = kx) \\
&= (1 - P_k)^{1-1} P_k + (1 - P_k)^{2-1} P_k + ... + (1 - P_k)^{kx-1} P_k \\
&= P_k((1 - P_k)^0 + (1 - P_k)^1 + ... + (1 - P_k)^{kx-1}) \\
&= P_k(1 + (1 - P_k)^1 + ... + (1 - P_k)^{kx} - (1 - P_k)^{kx}) \\
&= P_k(1 - (1 - P_k)^{kx} + \frac{(1 - P_k)(1 - (1 - P_k)^{kx})}{1 - (1 - P_k)}) \\
&= P_k(1 + \frac{-(1 - P_k)^{kx} + (1 - P_k)^{kx+1} + (1 - P_k) - (1 - P_k)^{kx+1}}{P_k}) \\
&= \frac{P_k}{P_k} * 1 - (1 - P_k)^{kx} \\
&= 1 - (1 - \frac{\omega}{k})^{kx} \\
&= 1 - ((1 - \frac{\omega}{k})^k)^x \\
&= 1 - e^{-\omega x} \\
&= \int_0^x \omega e^{-\omega x} dx \\
&= \int_{-\infty}^x \omega E^{-\omega x} dx \\
&= \int_{-\infty}^x f(x) dx.
\end{aligned}$$

Therefore, $f(x)$ is an exponential p.d.f., and we can say that $Z_k = \frac{N_k}{k}$ converges in distribution to Y as $k \to \infty$.

Question 3 .... *25 point*

A student is attempting a multiple choices exam. For each question there are 4 possible answers. He has a probability of 0.75 of knowing the correct answer. If he does not know the answer he chooses one answer uniformly and randomly. All questions and answers are independent.

To get a B he need to answer correctly 85% of the questions while to get an A he needs to answer correctly 95% of the questions.

To answer the questions below, you can use Matlab or R (or any other software) to compute the needed values c.d.f. of a Standard Normal r.v.. In case you do not have them available, this is an online calculator.

(a) (10 points) If the test contains 40 questions, use a normal approximation (CLT) to compute the probability p_B that the student will get at least a B and the probability p_A that the student will get a A.

In order to developed a normal distribution approximation, we first need to solve for a mean and standard deviation for this situation. First we will develop three probabilities,
p(he knows the answer) $= \frac{3}{4}$,
p(he doesn't know the answer but still gets it correct) $= \frac{1}{4}$
p(he gets the question correct) $= \frac{3}{4} + (\frac{1}{4})(\frac{1}{4}) = \frac{13}{16}$.
Therfore, the mean of this situation would simply be $\mu_x = \mathbb{E}(x) = n * p = 40 * \frac{13}{16} =$ 32.5 or 81.25%.. Similarly, the standard deviation will be $\sigma_x = \sqrt{\sigma_x^2} = \sqrt{var(x)} = \sqrt{n * p * (1 - p)} = \sqrt{40 * \frac{13}{16} * \frac{3}{16}} = \sqrt{\frac{195}{32}} \approx 2.47$.
Then, using a normal distribution, we know the probability that he gets at least a B is 0.27 or 27%, and the probability that he will get at least an A is 0.012 or 1.2%.

(b) (15 points) Let p_B be the probability that a student that knows 75% of the answers will get a B or more. If the teacher wants p_B to be less than 0.025, how many questions should there be on the exam.

Given that we want this probability to be less than 0.025, that means that we need the probability of getting a B to be 2 standard deviations away from the mean. Therefore, we can use the standard z-score formula for the normal distribution.

$$\begin{aligned} z &= \frac{x-\mu}{\sigma} \\ 2 &= \frac{0.85*n - n*\frac{13}{16}}{\sqrt{n*\frac{13}{16}*\frac{3}{16}}} \\ n &= 433.333. \end{aligned}$$

However, since there cannot be a third of a question, the teacher must have 434 questions on the examination.

Question 4 .... *20 point*

Let X be a continuous r.v. with uniform distribution in $[-1, 1]$ and Y be such that $\mathbb{P}(Y = 1) = \mathbb{P}(Y = -1) = 0.5$, X and Y independent. Consider the r.v. $Z = XY$.

(a) (10 points) Find the p.d.f. of Z.

We will first find the CDF of Z.

$$\begin{aligned}\mathbb{P}(Z \leq z) &= \mathbb{P}(XY \leq z)\\ &= \mathbb{P}(XY \leq z, Y = -1) + \mathbb{P}(XZ \leq z, Y = 1)\\ &= \mathbb{P}(X(-1) \leq z, Y = -1) + \mathbb{P}(X \leq z, Y = 1)\\ &= \mathbb{P}(X \geq -z)\mathbb{P}(Y = -1) + \mathbb{P}(X \leq z)\mathbb{P}(Y = 1)\\ &= (1 - \frac{1-z}{2})(\frac{1}{2}) + (\frac{z+1}{2})(\frac{1}{2})\\ &= \frac{1}{4}(z+1) + \frac{1}{4}(z+1)\\ &= \frac{z+1}{2} = F_Z(z).\end{aligned}$$

Then, PDF of Z will be $\frac{\partial F_Z(z)}{\partial Z} = \frac{1}{2}$. Then, $f_Z(z) = \frac{1}{2}$ for $-1 \leq Z \leq 1$.

(b) (10 points) Are Z and Y independent?

$\mathbb{E}(Z) = \int_{-1}^{1} z\frac{1}{2}dz = \frac{1}{2} * \frac{1}{2}[z^2]_{-1}^{1} = \frac{1}{4}(1-1) = 0$.
$\mathbb{E}(Y) = (-1)(\frac{1}{2}) + 1(\frac{1}{2}) = \frac{1}{2} - \frac{1}{2} = 0$
Now we will find the CDF and PDF of $ZY = M$

$$\begin{aligned}\mathbb{P}(M \leq m) &= \mathbb{P}(YZ \leq m)\\ &= \mathbb{P}(Z(-1) \leq m, Y = -1) + \mathbb{P}(Z \leq m, Y = 1)\\ &= \mathbb{P}(Z \geq -m)\mathbb{P}(Y = -1) + \mathbb{P}(Z \leq m)\mathbb{P}(Y = 1)\\ &= \frac{1}{2}(\frac{1+m}{2}) + \frac{1}{2}(\frac{m+1}{2}) = \frac{1}{2}(m+1) : -1 \leq m \leq 1.\end{aligned}$$

Then, $f_M(m) = \frac{d}{dm}(\frac{m+1}{2}) = \frac{1}{2} : -1 \leq m \leq 1$. Then the PDF of YZ is $f_M(m) = \frac{1}{2}$. $\mathbb{E}(YZ) = \mathbb{E}(M) = \int_{-1}^{1} \frac{1}{2}mdm = \frac{1}{4}[m^2]_{-1}^{1} = 0 -$. So as $\mathbb{E}(YZ) = \mathbb{E}(Y)\mathbb{E}(Z) = 0 * 0 = 0$, Y and Z are independent.

Question 5 .... *10 point*

Let X_1 and X_2 be two independent exponential r.v. with expected value λ. Find the joint p.d.f. of $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$ and the marginal p.d.f. of Y_2. (**Hint**: pay attention to the possible values of Y_1 and Y_2)

The expected value of an exponential r.v. is equal to the inverse of the parameter. That is to say that the parameter for both exponential r.v.s X_1 and X_2 is $\frac{1}{\lambda}$. That is to say that $f_{X_1}(x_1) = \frac{1}{\lambda} * e^{-\frac{1}{\lambda}x_1}$ and $f_{X_2}(x_2) = \frac{1}{\lambda} * e^{-\frac{1}{\lambda}x_2}$. Therefore, the p.d.f. of Y_1 is,

$$\begin{aligned}
f_{Y_1}(y_1) &= \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y_1 - x_1) dx_1 \\
&= \int_0^{y_1} \frac{1}{\lambda} e^{-\frac{1}{\lambda}x_1} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(y_1 - x_1)} dx_1 \\
&= \frac{1}{\lambda}\frac{1}{\lambda} e^{-\frac{1}{\lambda}y_1} \int_0^{y_1} e^{(\frac{1}{\lambda}-\frac{1}{\lambda})x_1} dx_1 \\
&= \frac{1}{\lambda^2} y_1 e^{-\frac{1}{\lambda}y_1}.
\end{aligned}$$

Similarly, the p.d.f. of Y_2 is,

$$\begin{aligned}
f_{Y_2}(y_2) &= \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(x_1 - y_2) dx_2 \\
&= \int_0^{\infty} \frac{1}{\lambda} e^{-\frac{1}{\lambda}x_1} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x_1 - y_2)} dx_2 \\
&= \frac{1}{\lambda} e^{\frac{1}{\lambda}y_2} \int_0^{\infty} \frac{1}{\lambda} e^{-2\frac{1}{\lambda}x} dx_2 \\
&= \frac{1}{2\lambda} e^{\frac{1}{\lambda}z}.
\end{aligned}$$

6. (15 points (bonus)) Let N_1, N_2 and N_3 be discrete random variables with joint probability mass function

$$p(n_1, n_2, n_3) = \mathbb{P}(N_1 = n_1 \,\&\, N_2 = n_2 \,\&\, N_3 = n_3) = \frac{3^{-N} N!}{n_1! n_2! n_3!}$$

if $n_1 + n_2 + n_3 = N$ and 0 otherwise.

Compute the marginal mass function p_{N_1} of N_1, that is

$$p_{N_1}(n_1) = \mathbb{P}(N_1 = n_1)$$

and the conditional mass function $p_{N_2,N_3|N_1}$ of N_2 and N_3 given N_1, that is

$$p_{N_2,N_3|N_1}(n_2, n_3|n_1) = \mathbb{P}(N_2 = n_2 \,\&\, N_3 = n_3 \,|\, N_1 = n_1)\,.$$

(**Hint**: you can answer the question without doing any computation. Think what situation is described by N_1, N_2 and N_3.)

$$\begin{aligned} p_{N_1} &= \sum_{n_1 \in N_2} \Big(\sum_{n_3 \in N_3} \Big(\frac{3^{-N} N!}{n_1! n_2! n_3!} \Big) \Big) \\ p_{N_2,N_3|N_1} &= \frac{p(n_1, n_2, n_3)}{p_{n_1}} \\ &= \frac{\frac{3^{-N} N!}{n_1! n_2! n_3!}}{\sum_{n_1 \in N_2} \big(\sum_{n_3 \in N_3} \big(\frac{3^{-N} N!}{n_1! n_2! n_3!} \big) \big)}. \end{aligned}$$

7. (15 points (bonus)) Let X_i, $i = 1, \ldots, N$ be independent and identically distributed continuous r.v. with median m, that is

$$\mathbb{P}(X_i \le m) = \frac{1}{2}\,.$$

Show that

$$\lim_{N\to\infty} \mathbb{P}\left(\min_{1\le i\le N}(X_i) \le m < \max_{1\le i\le N}(X_i)\right) = 1\,.$$