No books or notes allowed. No laptop or wireless devices allowed. Write clearly.

Name: ______________________________

Question:	1	2	3	4	5	Total
Points:	35	15	15	20	15	100
Score:						

Question 1 .... *35 point*

In a urn there are 1000 red balls and 1000 blue balls. You extract a ball at random and then flip a fair coin. If the result is Head you reinsert the ball you extracted. If the result is Tail you keep the ball. You repeat this procedure 20 times. Call N the r.v. that describes the total number of balls you have kept at the end of the 20 extractions and R the number of red balls among them.

(a) (10 points) Give the p.m.f. $f_N(n)$ of N and the conditional p.m.f. of R given $N = n$, that is $f_{R|N}(r|n)$

Solution: Clearly N is a binomial r.v. with $p = 0.5$ and 20 repetitions thus

$$f_N(n) = \binom{20}{n} 2^{-20}$$

On the other hand, if n is given, R is hypergeometric so that

$$f_{R|N}(r|n) = \frac{\binom{1000}{r}\binom{1000}{n-r}}{\binom{2000}{n}}$$

(b) (10 points) Write the j.p.m.f. of N and R and find an expression for the marginal $f_R(r)$ of R.

Solution: The j.p.m.f is easily found by mutiplying the results of point a):

$$f(n,r) = f_N(n) f_{R|N}(r|n) = 2^{-20} \frac{\binom{1000}{r}\binom{1000}{n-r}\binom{20}{n}}{\binom{2000}{n}}$$

while

$$f_R(r) = 2^{-20}\binom{1000}{r} \sum_{n>r} \frac{\binom{20}{n}\binom{1000}{n-r}}{\binom{2000}{n}}$$

(c) (15 points) Using that 20 << 2000 write an approximate marginal p.m.f for R and use it to compute $P(R = 5)$.(**Hint:** R can be seen as a binomial distribution ...)

Solution: Since 20 << 2000 we can assume that every extraction from the urn gives a blue ball with probability 0.5 and a red ball with probability 0.5. Thus the result of an extraction followed by the coin flip is a red ball with probability 0.25 and a blue ball or no ball with probability 0.75. We get that R is binomial with $p = 0.25$ and 20 repetiotions, that is:

$$P(R = r) \simeq \binom{20}{r} 0.25^r 0.75^{20-r}$$

and

$$P(R = 5) \simeq \binom{20}{5} 0.25^5 0.75^{15} = 0.202$$

Question 2 .... *15 point*

Let X_1 and X_2 be two independent r.v. with uniform distribution in $(0,1)$. Let $Y = \max(X_1, X_2)$. Find the p.d.f. of Y. (**Hint**: find $P(Y < y)$ for any given y.)

Solution: We have

$$\begin{aligned} F_Y(y) =& P(Y < y) = P(\max(X_1, X_2) < y) = P(X_1 < y \,\&\, X_2 < y) = \\ =& P(X_1 < y)P(X_2 < y) \end{aligned}$$

Clearly

$$P(X_1 < y) = P(X_2 < y) = \begin{cases} 0 & y < 0 \\ y & 0 < y < 1 \\ 1 & y > 1 \end{cases}$$

so that

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y^2 & 0 < y < 1 \\ 1 & y > 1 \end{cases}$$

Finally

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ 2y & 0 < y < 1 \\ 0 & y > 1 \end{cases}$$

Question 3 . . . . *15 point*

The following are six measurement of the waiting time between two successive nuclear decays in a Uranium sample:

2.379130 0.058303 0.587565 1.357028 1.263141 3.119415

You assume that these observations are independent and have a exponential distribution with parameter $\lambda = 1.8$. Build a q-q plot for these data. Remember you have first to compute the $100 * i/7$, $i = 1, \ldots, 6$, percentiles of the exponential distribution.

Solution: The c.d.f. of and exponential with parameter 1.8 is

$$F(x) = 1 - e^{-1.8*x}$$

We need the q_i such that

$$F(q_i) = i/7 \quad \Rightarrow \quad q_i = \frac{1}{1.8} \ln\left(\frac{7}{7-i}\right)$$

for $i = 1, 2, 3, 4, 5, 6$. We get for the

i	1	2	3	4	5	6
q_i	0.085639	0.186929	0.310898	0.470721	0.695979	1.081061

Calling y_i the order statistics we get

i	1	2	3	4	5	6
q_i	0.085639	0.186929	0.310898	0.470721	0.695979	1.081061
y_i	0.058303	0.587565	1.263141	1.357028	2.379130	3.119415

In Particular the assumption that $\lambda = 1.8$ is not verified. It looks more plausible that $\lambda \simeq 2$.

Question 4 . . . . *20 point*

Let X_1 and X_2 be two independent r.v. with uniform distribution in $(-1, 1)$. Let

$$Y_1 = X_1 + X_2$$
$$Y_2 = X_1 - X_2$$

(a) (10 points) Compute $E(Y_1)$, $E(Y_2)$, $V(Y_1)$, $V(Y_2)$ and $\mathrm{corr}(Y_1, Y_2)$.

Solution: We clearly have that

$$E(X_1) = E(X_2) = 0 \qquad V(X_1) = V(X_2) = \frac{1}{3}.$$

It follows that

$$E(Y_1) = E(X_1) + E(X_2) = 0, \qquad E(Y_2) = E(X_1) - E(X_2) = 0,$$
$$V(Y_1) = V(X_1) + V(X_2) = \frac{2}{3}, \qquad V(Y_2) = V(X_1) + V(X_2) = \frac{2}{3}.$$

Finally

$$E(Y_1Y_2) = E((X_1 + X_2)(X_1 - X_2)) = E(X_1^2) - E(X_2^2) = 0$$

so that

$$\mathrm{corr}(Y_1, Y_2) = 0.$$

(b) (10 points) Are Y_1 and Y_2 independent? Why?

Solution: No. Indeed if Y_1 is close to 2 it means that both X_1 and X_2 are close to 1 so that you know that Y_2 has to be close to 0.

Question 5 . . . . *15 point*

(a) (15 points) Let X be a normal r.v. with $E(X) = 2$ and $V(X) = 4$. Call $Y = 2X - 6$. Compute

$$P(-6 \leq Y \leq 2)$$

Solution:

$$P(-6 \leq Y \leq 2) = P(0 \leq X \leq 4) = P(-1 \leq Z \leq 1) = \Phi(1) - \Phi(-1)$$