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

Name: ______________________________

Question:	1	2	3	4	Total
Points:	32	36	22	10	100
Score:					

Question 1 . . . . *32 point*

The following numbers x_i, $i = 1, \ldots, 18$, represent a sample of size $n = 18$ from a given population.

0.9355	0.9169	0.4103	0.8936	0.0579	0.3529
0.8132	0.0099	0.1389	0.2028	0.1987	0.6038
0.2722	0.1988	0.0153	0.7468	0.4451	0.9318

(a) (12 points) Compute the sample median and fourth spread.

Solution:
After ordering the data you obtain

0.9355	0.9318	0.9169	0.8936	0.8132	0.7468
0.6038	0.4451	0.4103	0.3529	0.2722	0.2028
0.1988	0.1987	0.1389	0.0579	0.0153	0.0099

so that:

$$\tilde{x} = (0.4103 + 0.3529)/2 = 0.3816$$
$$lf = 0.1987 \qquad uf = 0.8132 \qquad fs = 0.8132 - 0.1987 = 0.6145$$

(b) (10 points) Knowing that $\sum_{i=1}^{18} x_i = 8.1442$ and , $\sum_{i=1}^{18} x_i^2 = 5.6743$ compute the sample mean and variance.

Solution:

$$\bar{x} = \frac{8.1442}{18} = 0.4525$$
$$s_X = \frac{1}{17}\left(5.6743 - \frac{8.1442 \cdot 8.1442}{18}\right) = 0.1170$$

(c) (10 points) Draw a box plot of the data. You do not need to check for outlier.

Solution:

Question 2 . . . . *36 point*

In Atlanta there are 2.000.000 families. Among them 40.000 do not report correctly their incomes. The IRS select a sample of 200 families and controlls their tax returns. Let X be the number of incorect reports among these 200.

(a) (12 points) Write a formula for the probability that $X = 4$.

Solution:

$$P(\{X = 4\}) = \frac{\binom{40.000}{4}\binom{1.960.000}{196}}{\binom{2.000.000}{200}}$$

(b) (12 points) Use a binomial approximation to compute the average and variance of X. Justify the approximation.

Solution: Since 200 is much smaller than 2.000.000 and 40.000 we can approximate X with a binomial r.v. with parameters $n = 200$ and $p = 0.02$. Thus we have:

$$E(X) \simeq np = 200 \cdot 0.02 = 4 \qquad V(X) \simeq np(1-p) = 200 \cdot 0.02 \cdot 0.98 = 3.92$$

(c) (12 points) Compute the probability that $X = 4$ using a Poisson approximation. Justify the approximation.(**Hint**; remeber that if X is a Poisson r.v. with parameter λ then $P(X = x) = \lambda^x e^{-\lambda}/x!$.)

Solution: Since n is large we can approximate the above binomial with a poissonian with parameter $\lambda = 200 \cdot 0.02 = 4$. We get

$$P(\{X = 4\}) = \frac{4^4 e^{-4}}{4!} = 0.1953$$

Question 3 .... *22 point*

In a bowl there are three balls numbered 0, 1 and 2. You randomly extract a ball. Then you put it back and randomly extract a ball again. Let X_1 be the result of the first extraction and X_2 the result of the second. Compute:

(a) (12 points) The pmf of $Y = X_1 - X_2$ and $Z = X_1 + X_2$.

Solution: There are 9 possible pair of balls (counting order). Each has the same probability to occur. Thus each pair has probability 1/9 to occur. The r.v. Y can take the 5 values: -2, -1, 0, 1, 2 while Z can take the values: 0, 1, 2, 3, 4. The following table explain the possibilities:

Y	pairs	Z	pairs
-2	(0,2)	0	(0,0)
-1	(0,1), (1,2)	1	(0,1), (1,0)
0	(0,0), (1,1), (2,2)	2	(2,0), (1,1), (0,2)
1	(1,0), (2,1)	3	(2,1), (1,2)
2	(2,0)	4	(2,2)

thus

$$p_Y(-2) = p_Y(2) = \frac{1}{9} \qquad p_Y(-1) = p_Y(1) = \frac{2}{9} \qquad p_Y(0) = \frac{1}{3}$$
$$p_Z(0) = p_Z(4) = \frac{1}{9} \qquad p_Z(1) = p_Z(3) = \frac{2}{9} \qquad p_Z(2) = \frac{1}{3}$$

(b) (10 points) The expected value and variance of Y and Z.

Solution: We have

$$E(Y) = (-2)\cdot\frac{1}{9} + (-1)\cdot\frac{2}{9} + 0\cdot\frac{1}{3} + 1\cdot\frac{2}{9} + 2\cdot\frac{1}{9} = 0$$
$$E(Z) = 0\cdot\frac{1}{9} + 1\cdot\frac{2}{9} + 2\cdot\frac{1}{3} + 3\cdot\frac{2}{9} + 4\cdot\frac{1}{9} = 2$$

while

$$E(Y^2) = 4\cdot\frac{1}{9} + 1\cdot\frac{2}{9} + 0\cdot\frac{1}{3} + 1\cdot\frac{2}{9} + 4\cdot\frac{1}{9} = \frac{4}{3}$$
$$E(Z^2) = 0\cdot\frac{1}{9} + 1\cdot\frac{2}{9} + 4\cdot\frac{1}{3} + 9\cdot\frac{2}{9} + 16\cdot\frac{1}{9} = \frac{48}{9}$$

so that

$$V(Y) = E(Y^2) - E(Y)^2 = \frac{4}{3}$$
$$V(Z) = E(Z^2) - E(Z)^2 = \frac{40}{9} - 4 = \frac{4}{3}$$

Question 4 .... *10 point*

In a bucket there are 3 balls. Some of them are red and the other are blue but you do not know how many are red. Call X the number of red balls. You only know that the p.m.f $p(x) = P(X = x)$ of X is:

$$p(0) = 0.2 \qquad p(1) = 0.2 \qquad p(2) = 0.3 \qquad p(3) = 0.3 \tag{1}$$

A ball is selected at random form the bucket and then reinserted. You see that the ball is red. Using this information, compute the probability that in the bucket there are 0, 1, 2 or 3 red balls. (**Hint**: you know that in the bucket there is at least one red ball. You have thus to compute conditional probabilities given this information. Use Bayes theorem:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \tag{2}$$

where B is an event and A_I is a family of mutually exclusive and exhaustive events)

Solution: Call A_i the event "there are i red balls in the bucket" and B the event "the selected ball was red". Eq. (1) gives the probability of the events A_i. On the other hand $P(B|A_i)$ is the probability of selecting a red ball when there are i red balls so that

$$P(B|A_i) = \frac{i}{3}$$

We can now opply Bayes theorem eq.(2). Observe that the denominator is equal for all i. So we first compute it:

$$\sum_j P(B|A_j)P(A_j) = 0.2\frac{1}{3} + 0.3\frac{2}{3} + 0.3 = 0.567$$

We thus get

$$P(A_0|B) = 0 \qquad P(A_1|B) = \frac{0.0667}{0.567} = 0.118$$
$$P(A_2|B) = \frac{0.2}{0.567} = 0.353 \qquad P(A_3|B) = \frac{0.3}{0.567} = 0.529 \tag{3}$$