Summer 2020

Name:____ **Total**______/100

MATH 4581: Statistics and Stochastic Processes

Test 1

Problem 1 Last year, five randomly selected students took a math aptitude test before they began their MATH 4581 course. In the table below, the x_i column shows scores on the aptitude test. Similarly, the y_i column shows the final grades.

Student	1	2	3	4	5	Sum	Mean
x_i	95	85	80	70	60	390	78
y_i	85	95	70	65	70	385	77

(a) [5 **pts**] Find the equation of regression line. Test at the 5% level of significance if the slope of regression line is 0.

Solution: The regression line is $y = 0.6438x + 26.78$ with $p = 0.1945 > 0.05$, implying the slope is not significantly different from zero, $s = 10.447$.

(b) [5 **pts**] Find the 95% confidence intervals for the mean and predicted values of y when $x = 75$. Show work.

Solution: First we find $\hat{y}(75) = 75.065$, $df = 5 - 2 = 3$ and the table t-value $t_{0.025,3} = 3.182$, so

$$CI_{mean} = 75.065 \pm 10.447 \cdot 3.182\sqrt{\frac{1}{5} + \frac{(75-78)^2}{730}} = 75.065 \pm 15.318,$$

$$CI_{predicted} = 75.065 \pm 10.447 \cdot 3.182\sqrt{1 + \frac{1}{5} + \frac{(75-78)^2}{730}} = 75.065 \pm 36.601.$$

Problem 2 Calcium is an essential mineral that regulates the heart, is important for blood clotting and for building healthy bones. The National Osteoporosis Foundation recommends a daily calcium intake of $1000 - 1200$ mg/day for adult men and women. A study is designed to test whether there is a difference in mean daily calcium intake in adults with normal bone density, adults with osteopenia (a low bone density which may lead to osteoporosis) and adults with osteoporosis. Adults 60 years of age with normal bone density, osteopenia and osteoporosis are selected at random from hospital records and invited to participate in the study. Each participant's daily calcium intake is measured based on reported food intake and supplements. The data are shown below.

Normal bone density	1200	1000	980	900	750
Osteopenia	1000	1100	800	700	500
Osteoporosis	350	900	400	890	650

(a) [5 **pts**] Use ANOVA at the 5% level of significance to decide if there is a statistically significant difference in mean calcium intake in patients with normal bone density as compared to patients with osteopenia and osteoporosis.

Solution: ANOVA$\longrightarrow$ p-value$= 0.1099 > 0.05$, so accept H_0 that there is a significant difference in mean calcium intake between the groups of patients.

(b) [10 **pts**] Test at the 5% level of significance to see if groups A='Normal bone density' and C='Osteoporosis' are different from group B='Osteopenia' using the contrast $C = \frac{3}{4}\mu_A + \frac{1}{4}\mu_C - \mu_B$.

Solution: We find $\bar{x}_A = 966, \bar{x}_B = 820$ and $\bar{x}_C = 638$, so the estimator $\widetilde{C} = \frac{3}{4}\bar{x}_A + \frac{1}{4}\bar{x}_C - \bar{x}_B = 64$. Since the number of patients in each group is 5 we have $n_a = n_b = n_c = 5$, which gives $S_C = \sqrt{MSE \cdot \frac{\left(\frac{1}{4}\right)^2+\left(\frac{3}{4}\right)^2+(-1)^2}{5}} = 224.91\sqrt{\frac{13}{40}} = 128.22$. Thus the experimental t-value is $t = \frac{\widetilde{C}}{S_C} = 0.499$ Next we find $df = 3{\cdot}5-3 = 12$, hence the critical value $t_c = t_{0.025,12} = 2.18$. As $0.499 < 2.18$ we accept the hypothesis.

Problem 3 [15 **pts**] Find the moment-generating function of the random variable Y having the distribution $f_Y(y) = 6(y-y^2)$ for $y \in [0,1]$.

Solution: $M_Y(t) = \mathbb{E}(e^{tY}) = \int_0^1 6(y-y^2)e^{ty}dy$. Using integration by parts we find

$\int_0^1 ye^{ty}dy = (\frac{1}{t}ye^{ty}\Big|_0^1 - \int_0^1 \frac{1}{t}e^{ty}dy) = \frac{e^t}{t} - \frac{e^{ty}}{t^2}\Big|_0^1 = \frac{(t-1)e^t+1}{t^2}$.

$\int_0^1 y^2e^{ty}dy = (\frac{1}{t}y^2e^{ty}\Big|_0^1 - \int_0^1 \frac{2}{t}ye^{ty}dy) = (\frac{1}{t}e^t - \int_0^1 \frac{2}{t}ye^{ty}dy) = (\frac{1}{t}e^t - (\frac{2}{t^2}ye^{ty}\Big|_0^1 - \int_0^1 \frac{2}{t^2}e^{ty}dy)) = (\frac{1}{t}e^t - \frac{2}{t^2}e^t + \frac{2}{t^3}e^t - \frac{2}{t^3}) = (\frac{1}{t} - \frac{2}{t^2} + \frac{2}{t^3})e^t - \frac{2}{t^3} = \frac{(t^2-2t+2)e^t-2}{t^3}$.

Finally, $M_Y(t) = 6(\frac{(t-1)e^t+1}{t^2} - \frac{(t^2-2t+2)e^t-2}{t^3}) = \frac{6(t-2)e^t+6t+12}{t^3}$

Problem 4 The transition matrix of a Markov chain is given below

$$M = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}.$$

(a) [5 **pts**] Find the probability that one moves from state 2 to state 3 in three steps. Show your work.

Solution:

$$M^3 = \begin{pmatrix} 0.25 & 0.375 & 0.375 \\ 0.5 & 0.25 & 0.375 \\ 0.375 & 0.375 & 0.25 \end{pmatrix}.$$

(b) [5 **pts**] Find the fixed vector with eigenvalue 1. Show your work.

Solution: The matrix equation $(x\ y\ 1-x-y)M = (x\ y\ 1-x-y)$ reads

$$\begin{cases} \frac{1}{2}(1-x) = x \\ \frac{1}{2}(1-y) = y \\ \frac{1}{2}(x+y) = 1-x-y. \end{cases} \tag{1}$$

Therefore the fixed vector is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Problem 5

Suppose you are given a Markov process with states $\{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\}$ and transition matrix

from/to	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{D}$
$\mathcal{A}$	1	0	0	0
$\mathcal{B}$	0.1	0.3	0.2	0.4
$\mathcal{C}$	0.3	0.5	0.1	0.1
$\mathcal{D}$	0	0	0	1

(a) [5 **pts**] List the absorbing states.

Solution: $\mathcal{A}, \mathcal{D}$.

(b) [5 **pts**] Rewrite the matrix in a canonical form and give Q.

Solution:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.1 & 0.3 & 0.2 & 0.4 \\ 0.3 & 0.5 & 0.1 & 0.1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{I} \begin{pmatrix} 0.3 & 0.5 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.2 & 0.4 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{II} \begin{pmatrix} 0.1 & 0.5 & 0.3 & 0.1 \\ 0.2 & 0.3 & 0.1 & 0.4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where transformation I exchanges the first and third row, followed by exchanging the first and third column via II. We find that $Q = \begin{pmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{pmatrix}$ (the first row and column correspond to state $\mathcal{C}$) and $R = \begin{pmatrix} 0.3 & 0.1 \\ 0.1 & 0.4 \end{pmatrix}$.

(c) [5 **pts**] Find the fundamental matrix N.

Solution: $N = (I-Q)^{-1} = \begin{pmatrix} 0.9 & -0.5 \\ -0.2 & 0.7 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{70}{53} & \frac{50}{53} \\ \frac{20}{53} & \frac{90}{53} \end{pmatrix}$.

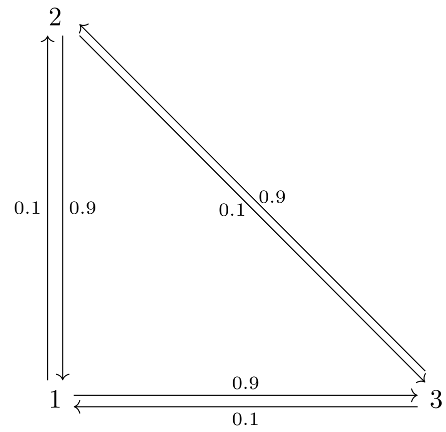
(d) [5 **pts**] Find the expected number of steps until absorption if we start in state $\mathcal{B}$.

Solution: $\frac{20}{53} + \frac{90}{53} = \frac{110}{53}$.

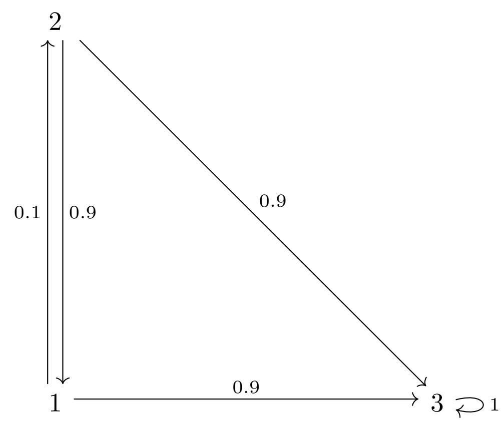
(e) [5 **pts**] Find the probability that we end in state $\mathcal{D}$ if we start in state $\mathcal{B}$.

Solution: $NR = \begin{pmatrix} \frac{26}{53} & \frac{27}{53} \\ \frac{15}{53} & \frac{38}{53} \end{pmatrix}$.

Problem 6 Consider the Markov chain with three states and transition probabilities as indicated on the graph below.



(a) [15 **pts**] Find m_{13}, the mean first passage time from the first to the third state.



Solution:

$$P' = \begin{pmatrix} 0 & 0.1 & 0.9 \\ 0.9 & 0 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$Q = \begin{pmatrix} 0 & 0.1 \\ 0.9 & 0 \end{pmatrix}.$$

$$N = (I - Q)^{-1} = \begin{pmatrix} \frac{100}{91} & \frac{10}{91} \\ \frac{90}{91} & \frac{100}{91} \end{pmatrix}.$$

$m_{13} = \frac{100}{91} + \frac{10}{91} = \frac{110}{91}$

(b) [10 **pts**] Find r_3, the mean recurrence time for the third state.

Solution:

$$m_{23} = \frac{90}{91} + \frac{100}{91} = \frac{190}{91}$$

$$r_3 = 1 + \frac{110}{91} \cdot 0.1 + \frac{190}{91} \cdot 0.9 = 3.$$