Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

FACULTY OF ENGINEERING, MATHEMATICS & SCIENCE

SCHOOL OF ENGINEERING

Electronic & Electrical Engineering

Engineering
Junior Sophister
Annual Examinations

Semester 1, 2018

Probability and Statistics (3E3)

Date: 13th December 2018 **Venue: RDS-SIM COURT** **Time: 14:00 – 16:00**

Anthony Quinn

Instructions to Candidates:
ANSWER QUESTION 1, and any THREE of the remaining five questions.

Question 1 is worth 30 marks *in toto.*
All remaining questions are worth equal marks (*i.e.* 70/3 marks each). The percentage division of marks within each question is indicated on the paper.
All symbols have their usual meaning.

Materials Permitted for this Examination:

Standard mathematical tables are available from the invigilators, if required;

Non-programmable calculators are permitted for this examination. Please indicate the make and model of your calculator on each answer book used.

Q.1 [Compulsory]

Answer ALL the following questions.

(a) Propositions S_1 and S_2 are (stochastically) independent. Each is (logically) sufficient for a third proposition, S_3. If $\Pr[S_1] = 0.4$ and $\Pr[S_3] = 0.8$, what is the maximal value of $\Pr[S_2]$? **[20 %]**

(b) Transistors can be in one of four modes, with probabilities 0.5, 0.3, 0.1 and 0.1, respectively. Five such transistors are probed. Write down an expression—which you do not need to evaluate—for the probability that at least three are operating in the 1st mode, while (at the same time) at most one is operating in the 4th mode. **[20 %]**

(c) A factory mass-produces a cheap device. It is known that about 1 in 10 are 'poor'. These devices are sequentially counted and tested until the next poor one is found. This number is recorded. What is the probability that this number is in the standard interval, adopting an appropriate probability model.

Hint: The geometric random variable, $N \sim \mathcal{G}e(p)$, $0 < p < 1$, has expected value, $m_N = \frac{1}{p}$, and variance, $\sigma_N^2 = \frac{1-p}{p^2}$. **[20 %]**

(d) A medical device is implanted in patients belonging to two groups, A and B, with twice as many A patients receiving the implant as B patients. The average lifetime of the implant is 8.5 years for patients in A, and 10.2 years for patients in B. When a particular patient presents with an implant, this is found still to be operational after 9.8 years. What is the probability that the patient belongs to group A? **[20 %]**

(e) The number of cars, n_i, $i = 1, \ldots, 6$, passing a point on a road in a 10-minute interval, is recorded at the same time of day and week, over 6 successive weeks, as follows:

n_i	18	16	22	13	18	21

Use these data to compute a parametric estimate of the probability that the number of cars next week, in 1 minute, will be no more than one (car) away from its mean.

[20 %]

Q.2

(a) Define conditional probability. Hence, state and prove Bayes' rule.

[30 %]

(b) Consider three propositions, S_1, S_2 and S_3, of probabilities p_1, p_2 and p_3, respectively. State the conditions under which these propositions are (stochastically) independent.

[20 %]

(c) An industrial plant produces components of either type 1 or type 2, in a ratio of 2:1.

- Quality assessments indicate that 6% of the type-1 components are defective, and 9% of the type-2 components are defective.
- The plant operates a quality control (QC) system which tests all the manufactured components. The QC system is able to detect 85% of defective type-1 components, and 75% of defective type-2 components. All perfect components are successfully passed by this system.

Specifying your probability model, consistent with the knowledge above, evaluate each of the following :

(i) the probability that a component produced by this plant is found to be defective by the QC system;

[20 %]

(ii) the probability that a component, which is found to be perfect by the QC system, is of type-2.

[30 %]

Q.3

(a) State and prove the binomial probability law.

[35 %]

(b) DNA molecules are formed from long sequences of bases, each base being of type G, C, A or T. In a particular DNA molecule, it is known that 25% of the bases are of type G, 15% are of type C, and 35% are of type A. An analysis of the sequencing of the bases in this molecule (using CRISPR) yields the following transition probability matrix from each base type to each base type (in the order above), for consecutive bases:

$$\mathrm{T} = \begin{bmatrix} 0.7 & 0.5 & 0.3 & 0 \\ 0.1 & 0.2 & 0.6 & 0 \\ 0.1 & 0.1 & 0.1 & 0.2 \\ 0.1 & 0.2 & 0 & 0.8 \end{bmatrix}.$$

(i) If the current base is of type A, what is the probability that the second-next base is also of type A?

[15 %]

(ii) What is the probability that five consecutive bases in this molecule are all of the same type?

[25 %]

(iii) Consider the bases at seven widely separated locations in this molecule. Write down an expression—which you do not have to evaluate—for the probability that exactly three of the bases are either of type G or C, while (at the same time) at least two of them are of type A.

[25 %]

Q.4

(a) State and interpret the Poisson probability law, specifying the conditions under which it is valid. **[30 %]**

(b) From (a), or otherwise, derive (i) the cumulative distribution function (cdf) for the time to the mth-next random Poisson point (RPP), $m \in \{1, 2, ...\}$; and (ii) the probability density function (pdf) for the time to the next RPP.

[25 %]

(c) Components in a particular large system are either in high demand or low demand, and these go into standby at an average rate of 2 per week and 1 per week, respectively. The system itself goes into standby when more than 3 components go into standby or if more than 2 high-demand components go into standby. Currently, none of the components is in standby. Stating any assumptions that you make, evaluate the probability that the system will not go into standby for more than one week.

[45 %]

Q.5

(a) A particle impinges randomly and uniformly on a square detector with vertices at (Cartesian) coordinates, $(1, 0)$, $(0, 1)$, $(2, 1)$ and $(1, 2)$ (arbitrary units). Let the coordinates of the particle be (X, Y).

(i) Plot the marginal distributions of X and Y, respectively, and evaluate the correlation coefficient, ρ_{XY}. Hence, argue that the coordinates are uncorrelated but not independent. **[30 %]**

(ii) Plot the regression function of Y, given $X = x$, including the standard interval in your plot. **[25 %]**

(b) Two closely spaced sensors on a bridge record displacements (in mm) caused by traversing vehicles. Future displacements recorded by sensors 1 and 2 are denoted by $D_1 = d_1$ and $D_2 = d_2$, respectively. The standard intervals of D_1 and D_2 are estimated as 1.9 ± 2.4 and -0.3 ± 1.1, respectively, and the correlation coefficient as -0.4. Justify the bivariate normal model for (D_1, D_2). Hence compute each of the following:

(i) the probability that $|D_1| < 0.5$; **[20 %]**

(ii) the probability that $|D_1| < 0.5$, if $D_2 = 0.5$. **[25 %]**

Q.6

(a) A particular type of medical device is fitted with two sensors. The lifetimes, in months, of these sensors, $(t_{1,i}, t_{2,i})$, $i = 1, 2, ..., 8$, are measured for each of 8 such devices under identical operating conditions, and recorded as follows:

$t_{1,i}$	5.8	4.8	21.9	6.6	7.2	30.9	36.2	11.3
$t_{2,i}$	1.2	0.4	6.0	19.8	19.0	13.6	1.7	13.7

Using these data, and stating any assumptions that you make, address each of the following, in respect of a future pair of measured sensor lifetimes, (T_1, T_2):

(i) Adopting a nonparametric model, estimate both $\Pr[T_2 > 10|T_1 > 20]$, and also $\Pr[T_1 > 20|T_2 > 10]$, as well as the correlation coefficient between T_1 and T_2.

[25 %]

(ii) Adopting an appropriate parametric model, estimate both $\Pr[T_1 > 20]$, and also $\Pr[T_2 > 10]$. Use your findings from (i) and (ii) to comment on possible dependence between T_1 and T_2. **[25 %]**

(b) A set of bipolar, real, scalar observations, $\mathrm{x}_n \equiv [x_1, \ldots, x_n]^T$, is recorded. The observations are known to be normally distributed with known variance, σ^2. The mean is either m_1 or m_2 (both known). Derive a statistical test of reduced form (i.e. avoiding unnecessary computations) to choose between m_1 and m_2, stating any assumptions that you make. **[50 %]**

A note on the bivariate normal distribution:

If

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}(\mathrm{m}, \Sigma), \quad \text{where} \quad \mathrm{m} = \begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix},$$

with

$$\sigma_{XY} = \rho_{XY}\sigma_X\sigma_Y,$$

then

$$f(x|Y = y) = \mathcal{N}(m_{X|y}, \sigma_{X|y}^2),$$

where

$$m_{X|y} = m_X + \rho_{XY}\frac{\sigma_X}{\sigma_Y}(y - m_Y),$$

$$\sigma_{X|y}^2 = \sigma_X^2\left(1 - \rho_{XY}^2\right).$$

Error Function Table

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

	Hundredths digit of x									
x	0	1	2	3	4	5	6	7	8	9
0.0	0.00000	0.01128	0.02256	0.03384	0.04511	0.05637	0.06762	0.07886	0.09008	0.10128
0.1	0.11246	0.12362	0.13476	0.14587	0.15695	0.16800	0.17901	0.18999	0.20094	0.21184
0.2	0.22270	0.23352	0.24430	0.25502	0.26570	0.27633	0.28690	0.29742	0.30788	0.31828
0.3	0.32863	0.33891	0.34913	0.35928	0.36936	0.37938	0.38933	0.39921	0.40901	0.41874
0.4	0.42839	0.43797	0.44747	0.45689	0.46623	0.47548	0.48466	0.49375	0.50275	0.51167
0.5	0.52050	0.52924	0.53790	0.54646	0.55494	0.56332	0.57162	0.57982	0.58792	0.59594
0.6	0.60386	0.61168	0.61941	0.62705	0.63459	0.64203	0.64938	0.65663	0.66378	0.67084
0.7	0.67780	0.68467	0.69143	0.69810	0.70468	0.71116	0.71754	0.72382	0.73001	0.73610
0.8	0.74210	0.74800	0.75381	0.75952	0.76514	0.77067	0.77610	0.78144	0.78669	0.79184
0.9	0.79691	0.80188	0.80677	0.81156	0.81627	0.82089	0.82542	0.82987	0.83423	0.83851
1.0	0.84270	0.84681	0.85084	0.85478	0.85865	0.86244	0.86614	0.86977	0.87333	0.87680
1.1	0.88021	0.88353	0.88679	0.88997	0.89308	0.89612	0.89910	0.90200	0.90484	0.90761
1.2	0.91031	0.91296	0.91553	0.91805	0.92051	0.92290	0.92524	0.92751	0.92973	0.93190
1.3	0.93401	0.93606	0.93807	0.94002	0.94191	0.94376	0.94556	0.94731	0.94902	0.95067
1.4	0.95229	0.95385	0.95538	0.95686	0.95830	0.95970	0.96105	0.96237	0.96365	0.96490
1.5	0.96611	0.96728	0.96841	0.96952	0.97059	0.97162	0.97263	0.97360	0.97455	0.97546
1.6	0.97635	0.97721	0.97804	0.97884	0.97962	0.98038	0.98110	0.98181	0.98249	0.98315
1.7	0.98379	0.98441	0.98500	0.98558	0.98613	0.98667	0.98719	0.98769	0.98817	0.98864
1.8	0.98909	0.98952	0.98994	0.99035	0.99074	0.99111	0.99147	0.99182	0.99216	0.99248
1.9	0.99279	0.99309	0.99338	0.99366	0.99392	0.99418	0.99443	0.99466	0.99489	0.99511
2.0	0.99532	0.99552	0.99572	0.99591	0.99609	0.99626	0.99642	0.99658	0.99673	0.99688
2.1	0.99702	0.99715	0.99728	0.99741	0.99753	0.99764	0.99775	0.99785	0.99795	0.99805
2.2	0.99814	0.99822	0.99831	0.99839	0.99846	0.99854	0.99861	0.99867	0.99874	0.99880
2.3	0.99886	0.99891	0.99897	0.99902	0.99906	0.99911	0.99915	0.99920	0.99924	0.99928
2.4	0.99931	0.99935	0.99938	0.99941	0.99944	0.99947	0.99950	0.99952	0.99955	0.99957
2.5	0.99959	0.99961	0.99963	0.99965	0.99967	0.99969	0.99971	0.99972	0.99974	0.99975
2.6	0.99976	0.99978	0.99979	0.99980	0.99981	0.99982	0.99983	0.99984	0.99985	0.99986
2.7	0.99987	0.99987	0.99988	0.99989	0.99989	0.99990	0.99991	0.99991	0.99992	0.99992
2.8	0.99992	0.99993	0.99993	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995	0.99996
2.9	0.99996	0.99996	0.99996	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998
3.0	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999	0.99999
3.1	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
3.2	0.99999	0.99999	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000