THE UNIVERSITY OF MELBOURNE

Student Number

Semester 1 Assessment, 2020

School of Mathematics and Statistics

MAST20006 Probability for Statistics

This exam consists of 28 pages (including this page)

Authorised materials: printed one-sided copy of the Exam or the Masked Exam made available earlier (or an offline electronic PDF reader), two double-sided A4 handwritten or typed sheets of notes, and blank A4 paper.

Calculators of any sort are NOT allowed.

Instructions to Students

- During exam writing time you may only interact with the device running the Zoom session with supervisor permission. The screen of any other device must be visible in Zoom from the start of the session.
- If you have a printer, print out the exam single-sided and hand write your solutions into the answer spaces.
- If you do not have a printer, or if your printer fails on the day of the exam,

  (a) download the exam paper to a second device (not running Zoom), disconnect it from the internet as soon as the paper is downloaded and read the paper on the second device;

  (b) write your answers on the Masked Exam PDF if you were able to print it single-sided before the exam day.

  If you do not have the Masked Exam PDF, write single-sided on blank sheets of paper.
- If you are unable to answer the whole question in the answer space provided then you can append additional handwritten solutions to the end of your exam submission. If you do this you MUST make a note in the correct answer space or page for the question, warning the marker that you have appended additional remarks at the end.
- Assemble all the exam pages (or template pages) in correct page number order and the correct way up, and add any extra pages with additional working at the end.
- Scan your exam submission to a single PDF file with a mobile phone or a scanner. Scan from directly above to avoid any excessive keystone effect. Check that all pages are clearly readable and cropped to the A4 borders of the original page. Poorly scanned submissions may be impossible to mark.
- Upload the PDF file via the Canvas Assignments menu and submit the PDF to the GradeScope tool by first selecting your PDF file and then clicking on Upload PDF.
- Confirm with your Zoom supervisor that you have GradeScope confirmation of submission before leaving Zoom supervision.
- You should attempt all questions.
- There is a table with the formulas for the main probability distributions and a table of standard normal distribution probabilities at the end of this question paper.
- There are 10 questions with marks as shown. The total number of marks available is 100.

Question 1 (7 marks)

It is known that 25% of all applicants to a language test have attended language coaching classes to prepare for the test. If an applicant has attended coaching classes, the probability is 0.8 that he or she will pass the language test, and if an applicant has not attended coaching classes, the probability is 0.5 that he or she will pass the test.

For a randomly selected applicant,

(a) State the probability that he/she has attended coaching classes.

(b) Find the probability that he/she passes the language test.

(c) Find the conditional probability that the selected applicant had attended coaching classes knowing that he/she has passed the language test.

Question 2 (14 marks)

A chocolate box contains 5 dark chocolates, 3 milk chocolates and 2 white chocolates.

(a) Suppose one chocolate is to be selected at random from the box. What is the value of p, the probability that a dark chocolate is to be selected?

(b) Suppose 3 chocolates are to be selected at random **with replacement**. Let X be the number of dark chocolates in the selection. What is the name and the associated parameter values of the distribution of X?

(c) Calculate $P(X = 1)$, the probability that there is one dark chocolate in the 3 selected.

(d) Suppose now that 25 chocolates are selected at random **with replacement** and denote by Y the number of dark chocolates in the selection.
Approximate the probability $P(Y > 10)$ by a normal probability using the continuity correction and the central limit theorem (the standard normal table is appended at the end of this exam paper).

(e) Justify why the normal approximation to the distribution of Y is more appropriate than a Poisson approximation here.

(f) Now suppose 4 chocolates are to be selected at random **without replacement**. Let W be the number of dark chocolates in the selection. What is the name and the associated parameter values of the distribution of W?

(g) Calculate $P(W = 3)$, the probability that there are three dark chocolates in the 4 selected.

Question 3 (11 marks)

In a company producing cheap bikes, the production process of bike brakes results in defective brakes 20% of the time. Assume that the quality status of any brake produced in this process is independent of the status of any other brake. A quality control inspector is to examine brakes one at a time to obtain three defective brakes. Let X be the number of bike brakes to examine until the inspector finds the first defective one, and let Y be the number of bike brakes to examine until she/he finds the third defective one.

(a) Name the probability distribution and specify the value of any parameter(s) for each of the two random variables X and Y.

(b) What is the probability that at least three brakes need to be examined until the inspector finds the first defective one?

(c) Determine the probability that the third defective brake is the 6th brake examined.

(d) *On average*, how many brakes need to be examined until the inspector finds the third defective one?

(e) Given that the first 2 brakes are not defective, what is the conditional probability that exactly 3 additional brakes need to be examined until the inspector finds the first defective one?

Question 4 (14 marks)

Let X be a continuous random variable with pdf (probability density function)

$$f(x) = c\,x^2 \quad \text{if } -1 < x < 2, \quad f(x) = 0 \text{ otherwise,}$$

for some constant c.

(a) Determine the value of c.

In the next subquestions, replace c by the value calculated in Part (a).

(b) Compute the probability $P(X > 1)$.

(c) Find the 25-th percentile of X (you are not required to write the answer in decimal form).

(d) Let $Y = |X|$ (absolute value of X) be a transformation of X.

(i) Is this transformation one-to-one? Find the support of Y.

(ii) Derive the cdf (cumulative distribution function) of Y.

(iii) Find the pdf of Y.

Question 5 (12 marks)

The mgf (moment-generating function) of a continuous random variable X is

$$M(t) = \frac{1}{(1-2t)^3}, \quad t < 1/2.$$

(a) State the probability distribution with name and value of any parameter(s) for the random variable X (if the distribution of X has more than one name, mention them all).

(b) How can X be expressed as a specific function of independent standard normal random variables?

(c) How can X be expressed as a specific function of independent exponential random variables?

(d) Find an expression (no numerical solution is required) for $P(X \leq \pi)$ by using the relationship between the distribution of X and a Poisson process.

(e) Calculate the mean $E(X)$ using the mgf.

(f) Calculate the second moment $E(X^2)$ using the mgf.

(g) Find the value of $\mathrm{Cov}(3X^2, X+1)$.

Question 6 (8 marks)

Three random variables X, Y and Z have the following mean and variance values: $E(X) = E(Y) = \sqrt{3}$, $E(Z) = 1, Var(X) = Var(Y) = 2$ and $Var(Z) = 9$. The three random variables are independent of each other. Let $W = XY + Z$.

(a) Compute $\mu = E(W)$.

(b) Compute $\sigma^2 = Var(W)$.

(c) Use Chebyshev's inequality $P(|W - \mu| < k\sigma) \geq 1 - (1/k^2)$ to find a lower bound for $P(-11 < W < 19)$.

Question 7 (10 marks)

Suppose X and Y are continuous random variables with joint pdf (probability density function)

$$f(x, y) = \begin{cases} \frac{1}{2}x + \frac{3}{2}y & \text{if } 0 \le x \le 1 \text{ and } 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal pdf of X.

(b) Find the conditional pdf of Y given $X = \frac{1}{4}$.

(c) Are X and Y independent? Why or why not?

(d) Compute the probability $P(Y > \frac{1}{2} \mid X = \frac{1}{4})$.

(e) Compute the probability $P(X^2 > Y)$.

Question 8 (9 marks)

Consider two random variables X and Y with the joint pdf (probability density function)

$$f(x,y) = \begin{cases} x+y, & \text{if } 0<x<1 \text{ and } 0<y<1, \\ 0 & \text{otherwise.} \end{cases}$$

Let $Z = XY$ and $U = X$ be a joint transformation of (X, Y).

(a) Find the support of (Z, U).

(b) Find the inverse transformation.

(c) Find the Jacobian of the inverse transformation.

(d) Find the joint pdf of (Z, U).

(e) Find the pdf of $Z = XY$ from the joint pdf of (Z, U).

Question 9 (10 marks)

Let $X_1, X_2, X_3, \ldots, X_n$ be a sample of independent random variables from a Poisson distribution with parameter λ.

(a) Find the mgf $M_{Y_n}(t)$ of the sum $Y_n = X_1 + X_2 + \ldots + X_n$.
Then name the distribution of Y_n and specify the associated parameter value(s).

(b) Find the mgf $M_{\bar{Y}_n}(t)$ of the sample mean $\bar{Y}_n = Y_n/n$.

(c) Find the limiting mgf, $\lim_{n\to\infty} M_{\bar{Y}_n}(t)$, using the result of Part (b). What distribution does the limiting mgf correspond to? What is the implication of this result?

Question 10 (5 marks)

Let N be the number of insurance claims received in one day by an insurance company. We assume that N has a Poisson distribution with mean 10. Let $X_1, X_2, \ldots$ be random variables representing claim amounts. We assume that N and $X_1, X_2, \ldots$ are independent, with $X_i \stackrel{d}{=} X$ (for all i) for some claim size X. Then $T = X_1 + X_2 + \ldots + X_N$ is the sum of a (random) number of independent random variables and represents the total amount claimed in one day.

Assume that X follows an exponential distribution with mean $\theta = 2$. Use the properties about the conditional mean and conditional variance to find

(a) The mean of T

(b) The variance of T.

End of Exam—Total Available Marks = 100

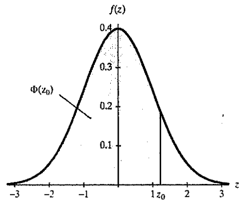
Table XII: Discrete Distributions

Probability Distribution and Parameter Values	Probability Mass Function	Moment-Generating Function	Mean $E(X)$	Variance $\text{Var}(X)$	Examples
Bernoulli $0 < p < 1$ $q = 1 - p$	$p^x q^{1-x}$, $x = 0, 1$	$q + pe^t$	p	pq	Experiment with two possible outcomes, say success and failure, $p = P(\text{success})$
Binomial $n = 1, 2, 3, \ldots$ $0 < p < 1$	$\binom{n}{x} p^x q^{n-x}$, $x = 0, 1, \ldots, n$	$(q + pe^t)^n$	np	npq	Number of successes in a sequence of n Bernoulli trials, $p = P(\text{success})$
Geometric $0 < p < 1$ $q = 1 - p$	$q^{x-1} p$, $x = 1, 2, \ldots$	$\dfrac{pe^t}{1 - qe^t}$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	The number of trials to obtain the first success in a sequence of Bernoulli trials
Hypergeometric $x \le n, x \le N_1$ $n - x \le N_2$ $N = N_1 + N_2$ $N_1 > 0$, $N_2 > 0$	$\dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$		$n\left(\dfrac{N_1}{N}\right)$	$n\left(\dfrac{N_1}{N}\right)\left(\dfrac{N_2}{N}\right)\left(\dfrac{N-n}{N-1}\right)$	Selecting r objects at random without replacement from a set composed of two types of objects
Negative Binomial $r = 1, 2, 3, \ldots$ $0 < p < 1$	$\binom{x-1}{r-1} p^r q^{x-r}$, $x = r, r+1, \ldots$	$\dfrac{(pe^t)^r}{(1 - qe^t)^r}$	$\dfrac{r}{p}$	$\dfrac{rq}{p^2}$	The number of trials to obtain the rth success in a sequence of Bernoulli trials
Poisson $0 < \lambda$	$\dfrac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, \ldots$	$e^{\lambda(e^t - 1)}$	λ	λ	Number of events occurring in a unit interval, events are occurring randomly at a mean rate of λ per unit interval
Uniform $m > 0$	$\dfrac{1}{m}$, $x = 1, 2, \ldots, m$		$\dfrac{m+1}{2}$	$\dfrac{m^2 - 1}{12}$	Select an integer randomly from $1, 2, \ldots, m$

Table XIII: Continuous Distributions

Probability Distribution and Parameter Values	Probability Density Function	Moment-Generating Function	Mean $E(X)$	Variance $\text{Var}(X)$	Examples
Beta $0 < \alpha$ $0 < \beta$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $0 < x < 1$		$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	$X = X_1/(X_1 + X_2)$, where X_1 and X_2 have independent gamma distributions with same θ
Chi-square $r = 1, 2, \ldots$	$\dfrac{x^{r/2-1} e^{-x/2}}{\Gamma(r/2) 2^{r/2}}$, $0 < x < \infty$	$\dfrac{1}{(1-2t)^{r/2}}$, $t < \dfrac{1}{2}$	r	$2r$	Gamma distribution, $\theta = 2$, $\alpha = r/2$; sum of squares of r independent $N(0, 1)$ random variables
Exponential $0 < \theta$	$\dfrac{1}{\theta} e^{-x/\theta}$, $0 \le x < \infty$	$\dfrac{1}{1 - \theta t}$, $t < \dfrac{1}{\theta}$	θ	θ^2	Waiting time to first arrival when observing a Poisson process with a mean rate of arrivals equal to $\lambda = 1/\theta$
Gamma $0 < \alpha$ $0 < \theta$	$\dfrac{x^{\alpha-1} e^{-x/\theta}}{\Gamma(\alpha)\theta^\alpha}$, $0 < x < \infty$	$\dfrac{1}{(1 - \theta t)^\alpha}$, $t < \dfrac{1}{\theta}$	$\alpha\theta$	$\alpha\theta^2$	Waiting time to αth arrival when observing a Poisson process with a mean rate of arrivals equal to $\lambda = 1/\theta$
Normal $-\infty < \mu < \infty$ $0 < \sigma$	$\dfrac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$, $-\infty < x < \infty$	$e^{\mu t + \sigma^2 t^2/2}$	μ	σ^2	Errors in measurements; heights of children; breaking strengths
Uniform $-\infty < a < b < \infty$	$\dfrac{1}{b-a}$, $a \le x \le b$	$\dfrac{e^{tb} - e^{ta}}{t(b-a)}$, $t \ne 0$; 1, $t = 0$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	Select a point at random from the interval $[a, b]$

Table Va: The Normal Distribution



$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

$$\Phi(-z) = 1 - \Phi(z)$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9545	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

α	0.400	0.300	0.200	0.100	0.050	0.025	0.020	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.054	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.240	2.326	2.576	2.807	3.291