Quiz 1: Probability and Statistics

5 marks

1. The CDF of a random variable X is defined as $F_X(x) = \mathbb{P}(\omega \in \Omega : X(w) \leq x) = \sum_{x \leq x_1} p_X(x)$ where p_X is the PMF. Prove that

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$$

2. A geometric random variable X with parameter p has PMF given by $p_X(k) = (1-p)^{k-1}p$. Derive the expression for its mean and variance.

3. Two cards are chosen from a standard deck of 52 cards. Suppose that you win 2 Rs for each heart selected, and lose 1 Rs for each spade selected. Other suits (clubs or diamonds) bring neither win nor loss. Let X denote your winnings. Determine the probability mass function $p_X(x)$.

4. For a random variable X with mean μ, its variance $Var(X)$ is defined as $E[(X-\mu)^2]$. Prove that $Var(aX+b) = a^2 Var(X)$ for arbitrary constants a and b.

10 marks

1. Let random variable X denote the outcome of a dice. Plot the Cumulative Distribution function (CDF) of X. Also find the mean and variance of X. Additionally prove that (prove! do not numerically verify. Start with either RHS or LHS and prove the other side.)

$$\sum_{x \in \{1,2,\ldots 6\}} x p_X(x) = 1 + \sum_{x \in \{1,2,\ldots 6\}} (1 - F_X(x))$$

(Hint: Write the CDF on the rhs in terms of PMF) The RHS is an alternative formula to get the expectation of non-negative random variables using the CDF.

Midsem: Probability and Statistics (50 Marks)

[Instruction: Please state reasons wherever applicable.]

5 marks

1. For a continuous non-negative random variable X prove that $E[X^2] = 2\int_x x\bar{F}_X(x)dx$ where $\bar{F}_X(x) = 1 - F_X(x)$.

2. Let X be a continuous random variable with distribution $F_X(\cdot)$ and density $f_X(x)$. Find the density and distribution for $Z = \sqrt{X}$.

3. Consider two exponential random variables X and Y with parameters λ_1 and λ_2 respectively. Consider $Z = min(X, Y)$ (min stands for minimum). Find the probability density and cumulative distribution of Z.

4. Let X and Y denote Gaussian random variables with mean μ_1 and μ_2 and standard deviation σ_1 and σ_2 respectively. Consider $Z = X + Y$. Using Moment generating functions, show that Z is also a Gaussian random variable, with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.

5. Let U_1 and U_2 be two independent Uniform random variables with support $[0, 1]$. Then find the cdf or pdf of $U_1 + U_2$.

6. If X and Y are independent random variables, prove that $Var(X+Y) = Var(X) + Var(Y)$. (Recall that $Var(X) = E[(X - E[X])^2])$

Please turn over for 10 marks questions

MA8.401: Probability and Statistics			Marks obtained ↓
Date: 26.11.2022,	Total questions: 10	Total points: 100	
Roll No: 202110107	Name:	Duration: 3 Hours	

Question:	1	2	3	4	5	6	7	8	9	10	Total
Points:	6	6	6	6	6	10	15	15	15	15	100
Score:											

Instructions:

1. **Explain all steps of your answer clearly.**
2. **Calculators are not allowed. Calculate approximately for two decimal digits.**
3. **For any doubt, seek help from invigilator.**

$\lambda e^{-\lambda x}$

1 Part 1 (30 Marks)

$\boxed{6}$ 1. Consider a sequence of i.i.d random variables $\{X_i, i \geq 1\}$ and let N be an independent random variable. Consider $T = \sum_{i=1}^{N} X_i$. Prove that $E[T] = E[N]E[X]$. Now let N denote a geometric random variable with parameter p and $X_i's$ are exponentially distributed with parameter λ (i.e. the pdf is given by $f_{X_i}(x) = \lambda e^{-\lambda x}$ for $x \geq 0$). Find $E[T]$.

$\boxed{6}$ 2. Consider two independent exponential random variables X and Y with parameters λ_1 and λ_2 (midsem) respectively. Consider $Z = \min(X, Y)$ and $W = \max(X, Y)$ (min stands for minimum and max stands for maximum). Find the cumulative distribution of Z and W.

$\boxed{6}$ 3. Let $Y = X^2$ where X is a standard normal random variable (Gaussian with zero mean and unit variance). Derive the CDF and pdf of Y. $\mu=0, \sigma=1$

$\boxed{6}$ 4. Derive the expression for the Moment generating function of the following random variables

- Uniform random variable with support $[a, b]$ (3 marks)
- Exponential random variable with parameter λ. (3 marks)

$\boxed{6}$ 5. Suppose V and W are both independent and exponentially distributed random variables with parameter μ. Find the CDF and pdf of Z where $Z = V - W$. (Rep)

2 Part 2 (70 Marks)

$L = pdf(x_1) \cdot pdf(x_2) \cdots pdf(x_n)$

MLE $\frac{dL}{d\theta} = 0$

6. For the following samples, find the maximum likelihood estimate for θ.

- $X_i \sim$ Exponential(θ) and we observed $(x_1, x_2, x_3, x_4) = (1.23, 3.32, 1.98, 2.12)$.

7. Let X_1, X_2, X_3, X_4, X_5 be a random sample from a $N(\mu, \overset{\sigma^2}{1})$ distribution, where μ is unknown. Suppose that we have observed the following values

$$5.45,\ 4.23,\ 7.22,\ 6.94,\ 5.98.$$

$\mu = \bar{X}$

We would like to decide between

$$H_0 : \mu = \mu_0 = 5,$$
$$H_1 : \mu \neq 5.$$

★. Define a test statistic to test the hypotheses and draw a conclusion assuming $\alpha = 0.05$

4. Find a 95% confidence interval around $\bar{X}$. Is μ_0 included in the interval? How does the exclusion of μ_0 in the interval relate to the hypotheses we are testing?

You have a coin and you would like to check whether it is fair or biased. More specifically, let θ be the probability of heads, $\theta = P(H)$. Suppose that you need to choose between the following hypotheses:

$$H_0: \text{ The coin is fair } \theta = \theta_0 = 1/2$$
$$H_1: \text{ The coin is not fair, } \theta > 1/2$$

We toss the coin 1000 times, and obtain 700 Heads.

1. Can we reject H_0 at significance level $\alpha = 0.05$?
2. Can we reject H_0 at significance level $\alpha = 0.01$?
3. What is the P-value?

Let X and Y be two jointly continuous random variables with joint PDF.

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

Find the MAP and the ML estimates of X given $Y = y$.

Suppose that the signal $X \sim N(0, \sigma_X^2)$ is transmitted over a communication channel. Assume that the received signal is given by

$$Y = X + W, \tag{1}$$

where $W \sim N(0, \sigma_W^2)$ is independent of X.

1. Find the MMSE estimator of X given Y, $(\hat{X}_M)$. **Hint:** Use the following

$$\hat{X}_M = E[X|Y] = \mu_X + \rho\sigma_X \frac{Y - \mu_Y}{\sigma_Y},$$

   where μ_X, μ_Y are means of X and Y, respectively, and σ_X^2, σ_Y^2 are variances of X and Y respectively.

2. Find the MSE of this estimator.

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997

Figure 1: The CDF $\Phi(z)$ values for $N(0,1)$. You may calculate $\Phi(-z) = 1 - \Phi(z)$. Also, you can calculate $\Phi^{-1}(z)$ or $\Phi^{-1}(-z)$ from this table. Some examples: $\Phi(2.63) = .99573$, $\Phi^{-1}(0.99960) \approx 3.35$.