No.

Date:

FIT 5197 Statistical Data Modelling
Assessment 1 : Aptitude Activity
Terminology, Probability and Integration

1. In $X = (x_1, \dots, x_n)$
   X is not necessarily ordered
   x_n is already observed, and not just a placeholder
   $\therefore x_n$ = n-th datapoint / observation

2. $P(X = a)$ can be shortened to $P(a)$
   $P(Y = \bar{b})$ can be shortened to $P(\bar{b})$

3. If $P(E) = P(E \mid F)$, knowing F doesn't give info about E
   E and F are independent

4. $P(E \mid F) = 1$ if E is a subset of F
   If F is a subset of E, $P(E \mid F) = 1$ still holds true.

5. For $P(x)$ to be a PMF, it must sum to 1

$$\sum_{x=1}^{x=2} \frac{1}{5} x^2 = \frac{1}{5}(1) + \frac{1}{5}(4)$$
$$= 1$$

   $\therefore$ For $P(x)$ to be a PMF, domain $x = 1, 2$

6. Cumulative Mass Function = summation or integral of Probability Mass Function

7. $\int_0^a (x^2 - 5x)\, dx = \frac{x^3}{3} - \frac{5x^2}{2}$

$$2x^3 = 15x^2$$
$$x = 7.5$$

Terminology: Probability and Integration

8. $f(x) = \begin{cases} x(x-1) & x\in[1,2] \\ 7-3x & x\in[2,\frac{7}{3}] \end{cases}$

$$\int_1^2 f(x) = \left[\frac{x^3}{3} - \frac{x^2}{2}\right]_1^2$$
$$= \left(\frac{8}{3} - \frac{4}{2}\right) - \left(\frac{1}{3} - \frac{1}{2}\right)$$
$$= (16 - 12 - 2 + 3)/6$$
$$= 5/6$$

$$\int_2^{7/3} f(x) = [7x - \tfrac{3}{2}x^2]_2^{7/3}$$
$$= \frac{49}{3} - \frac{3}{2}\left(\frac{49}{9}\right) - (14 - 6)$$
$$= 1/6$$

9. Following from Q 8.

$\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ → Constant for CDF $x \in [1, 2]$

$6 - 14 = -8$ → Constant for CDF $x \in [2, 7/3]$

Add in $\int_1^2 f(x) = \frac{5}{6}$

$-8 + \frac{5}{6} = -7\frac{1}{6}$

Inverse Functions and Quantile Functions

1. $f(x) = x^2 - 14$

$$y = x^2 - 14$$
$$y + 14 = x^2$$
$$x = \sqrt{y+14}$$

$$\therefore f^{-1}(x) = \sqrt{x+14}$$
$$f^{-1}(22) = \sqrt{36}$$
$$= 6$$

2. $F(x) = 1 - e^{-\lambda x} \quad x \geq 0$

Quantile function is inverse of CDF

$$y = 1 - e^{-\lambda x}$$
$$1 - y = e^{-\lambda x}$$
$$\ln(1-y) = -\lambda x$$
$$x = \frac{-\ln(1-y)}{\lambda}$$

$$Q(p) = x = \frac{-\ln(1-p)}{\lambda}$$

Counting and Probability

1. $P(\geq 2) = 1 - P(<2)$
$= 1 - (^{11}/_{12})(^{10}/_{12})(^{9}/_{12})(^{8}/_{12})$
$= 0.6181$

2. Arrange 4 books $= 4!$
Arrange the 4 book block and the other 6 books $= 7!$
Total permutations $= 4!\,7!$
$= 120\ 960$

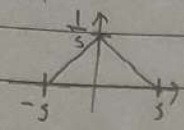
3. Get $[J, Q, K, A] = (4)(3)(2)(1)$ ways
Out of 52 cards $= (52)(51)(50)(49)$ total ways
$$P = \frac{4 \cdot 3 \cdot 2 \cdot 1}{52 \cdot 51 \cdot 50 \cdot 49}$$
There are 4 Js, Qs, Ks, As so
$$P = 4 \cdot 4 \cdot 4 \cdot 4 \cdot \frac{4 \cdot 3 \cdot 2 \cdot 1}{52 \cdot 51 \cdot 50 \cdot 49}$$

4. n balls, N boxes
$= N-1$ dividers, $n + N - 1$ items

Choose positions of balls
$$= \binom{n+N-1}{n}$$

Expectation, Chebychev and Absolute Values

1. $p(x \mid s) = \dfrac{s - |x|}{s^2} \qquad x \in [-s, s]$



$y = 0, \quad \dfrac{s - |x|}{s^2} = 0$

$s - x = 0 \qquad s + x = 0$

$x = s \qquad x = -s$

This is symmetric about 0
But we want $E[|X|]$, not $E[X]$, so:

$$E[|X|] = 2\int_0^s x \, \frac{s-x}{s^2} \, dx$$

$$= 2\int_0^s \frac{x}{s} - \frac{x^2}{s^2} \, dx$$

$$= 2\left[\frac{x^2}{2s} - \frac{x^3}{3s^2}\right]_0^s$$

$$= 2\left[\frac{s}{2} - \frac{s}{3}\right]$$

$$= \frac{1}{3}s$$

2. $P\left(\dfrac{|X - \mu|}{\sigma} \geq k\right) \leq \dfrac{1}{k^2}$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Magnitude of $(X - \mu)$ being more than $k\sigma$ has P within $\frac{1}{k^2}$
We want magnitude of $(X - \mu)$ less than $k\sigma$, so use complement rule to take inverse:

$$P(|X - \mu| < k\sigma) > 1 - \frac{1}{k^2}$$

Differentiation

1. $f(x) = x^n$ $\qquad f'(x) = n x^{n-1}$

$g(x) = e^x$ $\qquad g'(x) = e^x$

$h(x) = \ln x$ $\qquad h'(x) = 1/x$

2. $f(x) = (3x+5)e^{3x}$

$$f'(x) = (3x+5)(e^{3x})' + (3x+5)'(e^{3x})$$
$$= 3(3x+5)(e^{3x}) + (3)(e^{3x})$$
$$= (9x+18)e^{3x}$$

3. $\sigma(x) = \dfrac{1}{1+e^{-x}} = (1+e^{-x})^{-1}$

$$\sigma'(x) = -(1+e^{-x})^{-2}(-e^{-x})$$
$$= e^{-x}(1+e^{-x})^{-2}$$

$$\sigma(x) = \frac{1}{1+e^{-x}},$$
$$1+e^{-x} = \frac{1}{\sigma(x)}$$
$$e^{-x} = \frac{1}{\sigma(x)} - 1$$
$$= \frac{1-\sigma(x)}{\sigma(x)}$$

$$\sigma'(x) = e^{-x}(1+e^{-x})^{-2}$$
$$= \frac{1-\sigma(x)}{\sigma(x)}\left(\frac{1}{\sigma(x)}\right)^{-2}$$
$$= \frac{1-\sigma(x)}{\sigma(x)} . (\sigma(x))^2$$
$$= \sigma(x)[1-\sigma(x)]$$

Functions, Operations and Stationary Points

1. $p^{-qy_1}(1-p) \cdot p^{-qy_2}(1-p) \ldots$

$$= p^{-qy_1 - qy_2 \ldots}(1-p)^n$$

$$= p^{\Sigma -qy_i}(1-p)^n$$

2. $\frac{d}{dp}\left[-\log \prod_{i=1}^{n} p^{-qy_i}(1-p) \right]$

$$= \frac{d}{dp}\left[-\log \left(p^{\Sigma -qy_i}\right)(1-p)^n \right] \quad \text{(From Q1)}$$

$$= \frac{d}{dp}\left[(\Sigma -qy_i)\cdot(-\log p) + (n)(-\log(1-p)) \right]$$

$$= \frac{d}{dp} \Sigma qy_i \log p - n \log(1-p)$$

$$= \frac{\Sigma qy_i}{p} + \frac{n}{1-p}$$

3. From Q2, stationary point is:

$$\frac{\Sigma qy_i}{p} + \frac{n}{1-p} = 0$$

$$(1-p)\, q \Sigma y_i + np = 0$$

$$np + q\Sigma y_i - pq\Sigma y_i = 0$$

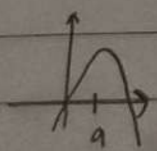
Solve for p:

$$p\,(n - q\Sigma y_i) = -q\Sigma y_i$$

$$p = \frac{-q\Sigma y_i}{n - q\Sigma y_i}$$

$$= \frac{q\Sigma y_i}{q\Sigma y_i - n}$$

4. For $x = a$ stationary point to be a maximum



$x < a$ is increasing $\therefore f'(x) > 0$

$x > a$ is decreasing $\therefore f'(x) < 0$