1	2	3	4	Total

Name: Answers
Number:

KOM505E - Probability Theory and Stochastic Processes
Midterm #1

Nov. 1, 2018

Rules:
- Closed book & notes.
- Write all answers within the frame given below the question.
- Duration: 90 min.

1. Assume there are 30 students (including you) in KOM505E course. For this question, ignore leap years (assume all years have 365 days). Just write the formula. There is no need to compute an exact value.

   (a) write the probability that **exactly** three students (in the course) have the same birthday with you.

   k = number of students that have the same birthday with you.
   N = 29 (students in KOM505E excluding you) $p = 1/365$

   $$P(k=3) = \binom{29}{3} p^3 (1-p)^{29-3}$$

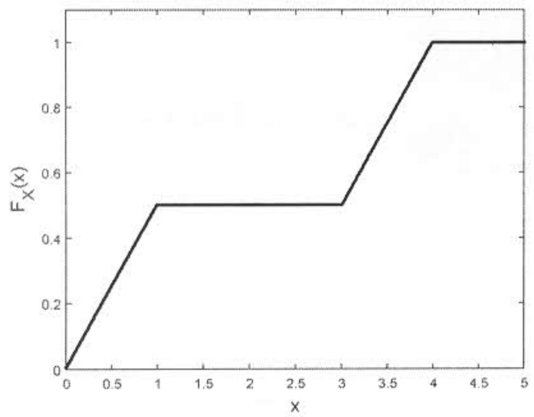
   (b) write the probability that **at least** three students (in the course) have the same birthday with you.

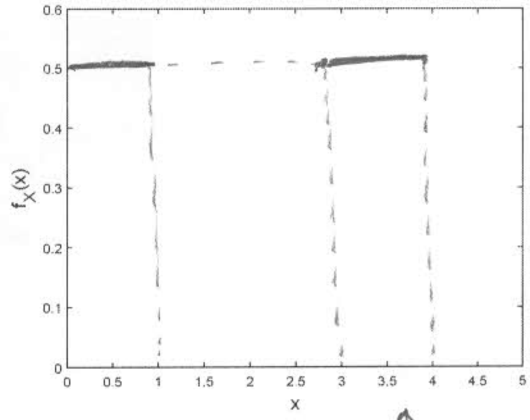
   $$P(k \geq 3) = 1 - \underbrace{P(k<3)}_{[P(k=0)+P(k=1)+P(k=2)]}$$
   $$= 1 - \binom{29}{0}(1-p)^{29} - \binom{29}{1} p \cdot (1-p)^{28} - \binom{29}{2} p^2 (1-p)^{27}$$

   (c) write the probability that **at most** three students (in the course) have the same birthday with you.

   $$P(k \leq 3) = P(k=0) + P(k=1) + P(k=2) + P(k=3)$$
   $$= \underbrace{P(k<3)}_{\text{computed in part (b)}} + \underbrace{P(k=3)}_{\text{computed in part (a)}}$$

2. (20 pts) Consider a continous random variable X. The cumulative distribution function, $F_X(x)$ of this rv is given below at the left figure.





(a) Draw the probability density function (pdf) of X on the right figure given above.

(b) Find the mean μ_X and variance σ_X^2 of X.

$$\mu_X = E(X) = \int_0^1 x \underbrace{f_X(x)}_{0.5} dx + \int_3^4 x \underbrace{f_X(x)}_{0.5} dx$$

$$= \frac{x^2}{4}\Big|_0^1 + \frac{x^2}{4}\Big|_3^4 = \frac{1}{4} + \frac{16-9}{4} = \frac{8}{4} = 2$$

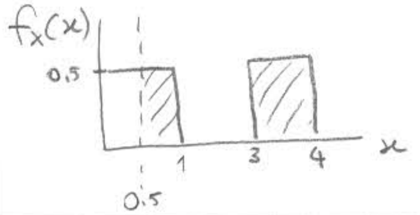
$$E(x^2) = \int_0^1 x^2 . 0.5 \, dx + \int_3^4 x^2 . 0.5 \, dx = \frac{x^3}{6}\Big|_0^1 + \frac{x^3}{6}\Big|_3^4$$

$$= \frac{1}{6} + \frac{64-27}{6} = \frac{38}{6} = \frac{19}{3}$$

$$\sigma_X^2 = E(x^2) - \mu_X^2 = \frac{19}{3} - 4 = \frac{7}{3}$$

(c) Find the following probabilities:

i. $P(0.5 < X)$



$$P(0.5 < x) = 1 - \underbrace{P(x < 0.5)}_{F_x(0.5) = 0.25} = 0.75$$

ii. $P(0.5 < X < 3.5)$

$$P(0.5 < x < 3.5) = F_x(3.5) - F_x(0.5)$$

$$= 0.75 - 0.25$$

$$= 0.5$$

1	2	3	4	Total

Name:
Number: Answer

3. Let X be the life expectancy of a cigarette smoker and Y be the life expectancy of a non-smoker. Assume both X and Y have normal (Gaussian) distribution. $X \sim N(70, 9)$ and $Y \sim N(77, 25)$.

For Gaussian distribution, the pdf and cdf are as follows:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right\}$$

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{x} \exp\left\{-\frac{(t-\mu_X)^2}{2\sigma_X^2}\right\} dt$$

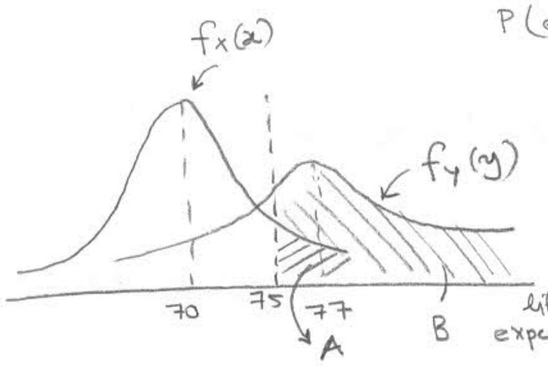
(a) What is the probability of a person that dies between ages of 73–75 be a smoker? Assume the smoking probability is 60% in the population. Write the probability in terms of $f_X(x)$ and/or $F_X(x)$. There is no need to compute a value.

$$P(\text{smoker} \mid 73 < x < 75) = \frac{P(73 < x < 75 \mid \text{smoker}) \cdot P(\text{smoker})}{P(73 < x < 75)}$$

$$= \frac{(F_X(75) - F_X(73))\, 0.6}{(F_X(75) - F_X(73))\, 0.6 + (F_Y(75) - F_Y(73))\, 0.4}$$

(b) Doctors assume that who lived more than 75 years is a non-smoker. What is the probability of doctors being wrong?

i. Draw the pdf of X and Y on a single graph, and shade area that corresponds to the error probability.



$$P(\text{error}) = P(\text{smoker} \mid 75 < x)$$

$$= \frac{P(75 < x \mid \text{smoker})\, 0.6}{P(75 < x)}$$

$$A \triangleq 1 - F_X(75)$$

$$B \triangleq 1 - F_Y(75)$$

$$P(\text{error}) = \frac{A \times 0.6}{A \times 0.6 + B \times 0.4}$$

ii. Write the error probability in terms of $f_X(x)$ and/or $F_X(x)$. There is no need to compute a value.

$$P(\text{error}) = \frac{(1 - F_X(75))\, 0.6}{(1 - F_X(75))\, 0.6 + (1 - F_Y(75))\, 0.4}$$

4. Let X be a continous random variable that has Gaussian distribution whose expected value and variance are $\mu_X = 5$ and $\sigma_X^2 = 36$ respectively. Find a function that convert X to a new continous random variable Y that has exponential distribution with $\lambda = 1$.

For Gaussian distribution, the pdf and cdf are as follows:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right\}$$

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{x} \exp\left\{-\frac{(t-\mu_X)^2}{2\sigma_X^2}\right\} dt$$

For exponential distribution, pdf is $f_Y(y) = \lambda \exp\{-\lambda y\}$ when $y \geq 0$ and $f_Y(y) = 0$ when $y < 0$. The cdf is $F_Y(y) = 1 - \exp\{-\lambda y\}$ when $y \geq 0$.

$$X \xrightarrow{g_1(X)} Z \xrightarrow{g_2(Z)} Y \quad \text{where } Z \sim \mathrm{Uni}(0,1)$$

$$Z = g_1(X) = F_X(x) \quad \text{and} \quad Y = g_2(Z) = F_Y^{-1}(Z)$$
$$= -\ln(1-Z)$$

then

$$Y = -\ln(1 - F_X(x))$$

Alternatively

$$F_X(x) = F_Y(y)$$
$$F_X(x) = 1 - e^{-y} \Rightarrow e^{-y} = 1 - F_X(x)$$
$$-y = \ln(1 - F_X(x))$$
$$y = -\ln(1 - F_X(x))$$