Note: Answers to all versions of a question are provided here.

Question 1 (6x2.5 points)

<u>Special Instructions for this question:</u> You **DO NOT** have to show work for (a) through (e). Just fill in the blanks. Any scratch work will not be evaluated, only the answer on the line matters!

a) The probability mass function of a discrete random variable *X* is defined as *p*(*x*) = *kx* for *x* = 1,2,3,4, then the value of *k* is __________.

ANSWER: k=1/10 = 0.10 [Alternate Version: For the support set x=1,2,3,4,5; **k=1/15**]

b) The probability mass function *p*(*x*) of a discrete random variable *X* is *p*(*0*) = .15, *p*(1) = .30, *p*(2) = .20, *p*(3) = .10, and *p*(4) = .25, then the value of the cumulative distribution function *F*(*x*) at X = 2 is __________.

ANSWER: Sum of probabilities at 0,1,2 = 0.65 [Alternate Version: *p*(*0*) = .15, *p*(1) = .30, *p*(2) = .30, *p*(3) = .10, and *p*(4) = .15, The sum of probabilities at 0,1,2 is **0.75**]

c) Let *X* be a discrete random variable with $E(X^2)$ = 19.75 and *V*(*X*) = 16.3275, then *E*(*X*) = __________.

ANSWER: Using V(X) = E(X^2) - [E(X)]^2, we get [E(X)]^2 =19.75-16.3275 = 3.4225. So, E(X) = ± 1.85

d) If the random variable *X* has a Poisson distribution with parameter $\lambda = 4$, then $E(X^2)$ is __________.

ANSWER: E(X^2) = 4^2+4 = 20, since for Poisson, E(X) = Var(X) = $\lambda = 4$.
ALTERNATE VERSION: For $\lambda = 5$ the answer is **E(X^2) = 5^2+5= 30**

e) If the assembly time for a product is uniformly distributed between 15 to 20 minutes, then the probability of assembling the product between 16 to 18 minutes is __________.

ANSWER: The pdf is a rectangle with base (15, 20) and height = 1/5 = 0.2. The area under the constant function in the interval (16, 18) equals 2x0.2=0.4.

f) If the *pdf* of a continuous random variable *X* is

$$f(x) = \begin{cases} .5x & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

, then $P(1 \le x \le 1.5)$ is **<u>0.3125</u>**.

Alternate Version: For $f(x) = \begin{cases} \frac{1}{8}x & 0 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$ the probability between 2 and3 is also **0.3125** .

Question 2

(25 points)

The probability density function of the time X (in minutes) that a flight from Columbus to Detroit arrives earlier or later than its scheduled arrival is given by

$$f(x) = \begin{cases} k(5-x)(5+x) & \text{for } -5 \le x \le 5, \\ 0 & otherwise. \end{cases}$$

Note: The negative values of x indicate flight arriving early, while positive values of x indicate flight arriving late.

(a) [5 points] Find the value of the constant k, so that $f(x)$ is a valid pdf.

Answer: Note that $f(x) = f(-x)$, therefore the above *pdf* is symmetric around zero. Now

$$k\int_{-5}^{5}(25-x^2)\,dx = 2k\int_{0}^{5}(25-x^2)\,dx = 2k\,(25x-\frac{x^3}{3})\Big|_0^5$$

$$= 2k[(25*5-\frac{1}{3}5^3)-0] = 2k*5^3*\frac{2}{3} = \frac{500}{3}k$$

$$\text{Now } \frac{500}{3}k = 1 \Rightarrow k = 3/500.$$

(b) [2 points] Find E(X) without calculating the actual integral $E[X] = \int x\,f(x)\,dx.$

Answer: Since the *pdf* is symmetric around 0, *E[X]* = 0.

(c) [5 points] Find the probability that one of these flights will arrive between 1 to 3 minutes earlier than its scheduled arrival.

Answer: The event that the flight will arrive between 1 to 3 minutes earlier than its scheduled arrival can be written as $A = \{-3 \le X \le -1\}$. Now

$$P(A) = \int_{-3}^{-1}\frac{3}{500}(25-x^2)dx = \frac{3}{500}(25x-\frac{x^3}{3})\Big|_{-3}^{-1}$$

$$= \frac{3}{500}[-(25-\frac{1}{3})+(25*3-9)] = \frac{124}{500} = .248$$

(d) [3 points] Find the probability that one of these flights will arrive between -3 to 3 minutes of its scheduled arrival.

ANSWER: Since the pdf is symmetric around zero, the probability of this event is twice the probability of the event {*0 < X < 3}. However, using the expression of the integral in part (a),*

$$P(-3 \le X \le 3) = 2k\int_0^3 (25 - x^2)\, dx = 2k\,(25x - \frac{x^3}{3})\Big|_0^3$$
$$= 2k[(25*3 - \frac{1}{3}3^3) - 0] = 2k*66 = 2*\frac{3}{500}*66 = .792$$

Note: For the next two parts of this problem, suppose that the flight durations for each flight from Columbus to Detroit are independently distributed with the *pdf* in part (a). A flight is considered to have an ON-TIME arrival if it arrives within 3 minutes of its scheduled arrival time. The probability of an on-time arrival is given in part (d) above. [Note: If you are not sure about your answer to part (d), use an approximate value of 0.8 for this probability.]

(e) **(5 Points)** Let the random variable Y denote the number of on-time arrivals among the next 10 flights. What is the distribution of Y? Find E(Y) and standard deviation of Y.

Answer: Since the arrival times are independent random variables, the distribution of Y is Binomial (n=10, p =0.792). You can use p=0.8.

Therefore, E(Y) = n p =10 x 0.792 =7.92 or ~8.0,
Var (Y) = *n p(1-p)*=10*.8 *.2 = 1.6 , or $\sigma_Y = 1.265$

(f) **(5 Points)** Let the random variable Z denote the number of on-time arrivals before the **first flight** that was NOT an on-time arrival. What is the pmf of Z?

The distribution of Z is a Negative Binomial, nb(x; r=1, P(S)=0.2).
Explanation- In the counting for Z, the ON-Time Arrival is labeled as Failure, and NOT a On-Time is labeled as Success. We are counting the **# of failures** before the **first** success. The Probability of Success is 1-*P*(Failure) = 1-0.8 = 0.2, So the *pmf* of Z is given by
p(z)= $P(Z = z) = nb(z;1,.2) = (0.2)(0.8)^z, z = 0,1,\cdots$
This is also known as the geometric distribution.

Question 3 [Alternate Version Question 5] **(10 points)**

Radioactive substances emit alpha particles during their decay process. The number of such particles emitted, as measured by a counter, follows a Poisson process with the rate $\boldsymbol{\alpha = 0.5}$ **per 10 seconds**. Suppose the half-life period of this substance is 2 minutes.

(a) (5 points) Find the probability that **at most 5** alpha particles are emitted **during one half-life period**.

Answer: one half-life = 2 minutes = 120/10 (10 sec periods)= 12.
Therefore X = # of alpha-particles in this duration ~ Poisson($\lambda = 12 * 0.5 = 6$).

P(X is at most 5) = $F(5;\lambda = 6) = \sum_{x=0}^{5} e^{-6} \frac{6^x}{x!} = .446$

Note: You didn't have to calculate the final number here.

(b) **(5 points)** What is the **distribution** of the number of alpha particles emitted during a **12 minute period**?

Answer: 12 minutes period = 12 x 6 =72 (10 seconds periods). Therefore, the distribution of the # of particles in a 12 minutes period is Poisson with $\lambda = 72 * 0.5 = 36$.

Question 4 [Alternate version: Q 3] **(15 points)**

An insurance company offers its policyholders a number of different payment options. For a randomly selected policyholder, let X = the number of months between successive payments. The *cdf* of X is as follows:

$$(i)F(x)=\begin{cases}0 & x<1\\ .20 & 1\le x<3\\ .40 & 3\le x<4\\ .45 & 4\le x<6\\ .70 & 6\le x<12\\ 1 & 12\le x\end{cases} \quad \text{Alternate (ii) } F(x)=\begin{cases}0 & x<1\\ .20 & 1\le x<3\\ .30 & 3\le x<4\\ .45 & 4\le x<6\\ .70 & 6\le x<12\\ 1 & 12\le x\end{cases}$$

a. **(6 points)** Find the pmf of X?

Answer: Possible *X* values are those values at which *F*(x) jumps, and the probability of any particular value is the size of the jump at that value. Thus we have :

x	1	3	4	6	12
(i)P(*x*)	.20	.20	.05	.25	.30
(ii)P(*x*)	.20	.10	.15	.25	.30

b. **(4 points)** Using just the cdf, compute $P(3\le X\le 6)$.

Answer:
$$(i)P(3\le X\le 6)=F(6)-F(3-)=.70-.20=.50$$
$$(ii)P(3\le X\le 6)=F(6)-F(3-)=.70-.20=.50$$

c. **(5 points) Find *E(X)* and *E(X²)*.**

Answer:

$$(i)E(X) = 1*.2+3*.2+4*.05+6*.25+12*.3 = 6.1$$
$$E(X^2) = 1*.2+9*.2+16*.05+36*.25+144*.3 = 55.0$$
$$(ii)E(X) = 1*.2+3*.1+4*.15+6*.25+12*.3 = 6.2$$
$$E(X^2) = 1*.2+9*.1+16*.15+36*.25+144*.3 = 55.7$$

Question 5 [Alternate Ver Q 4 in BOLD #] **(15 points)**

(a) **(8 points)** A geologist has collected 10 **[15]** specimens of basaltic rock and 10 **[15]** specimens of granite. The geologist instructs a laboratory assistant to randomly select 15 **[20]** of the specimens for analysis.

I. **(5 points)** What is the set of all possible values of the random variable *X* = Number of granite specimens selected for analysis? Find *P(X=7)* and *P(X=8)*.

ANSWER: Possible values of *X* are 5,6,7,8,9,10. (In order to have less than 5 of the granite, there would have to be more than 10 of the basaltic).

$$(i) P(X=7) = h(7;15,10,20) = \frac{\binom{10}{7}\binom{10}{8}}{\binom{20}{15}} = 0.3483.$$ Note: Becuase of symmetry, P(X=8) has the same value.

Following the same pattern for the other values, we arrive at the pmf, in table form below. But you didn't need to calculate all of them.

x	5	6	7	8	9	10
P(*x*)	.0163	.1354	.3483	.3483	.1354	.0163

Alternate Version: Possible values of *X* are 5,6,7,8,9,10,⋯, 15. (In order to have less than 5 of the granite, there would have to be more than 15 of the basaltic).

$$(ii) P(X=x) = h(x;20,15,30) = \frac{\binom{15}{x}\binom{15}{20-x}}{\binom{30}{20}}, x = 5,\ldots,15.$$

Note: Because of symmetry, P(X=9) = P(X=11) has the same value.

Calculating the above values for *x* = 9, 10, 11, we get

x	9	10	11
P(*x*)	.2274	.3001	.2274

II. **(3 points)** What is the probability that the number of granite specimens selected for analysis is within 1 standard deviation of its mean value?

Answer: E(X) = *n* x (*M* / *N*) = 15 (10 / 20) = 7.5;
V(X) = (5/19) 15 (10 / 20) [1 – (10/20)] = .98684; σ_x = .9934
$\mu \pm \sigma =$ 7.5 $\pm$ .9934 = (6.5066, 8.4934), so we get

P(X = 7) + P(X = 8) = .3483 + .3483 = .6966

ALTERNATE VERSION: Find E(X) and standard deviation of X?

Answer: E(X) = *n* x (*M* / *N*) = 20 (15 / 30) = 10;
V(X) = (10/29) 20 (15 / 30) [1 – (15/30)] = 1.7241; σ_x = 1.3130

(b) (7 points) Assume that 1 in 200 people carry the defective gene that causes inherited colon cancer.

I. (2 points) In a sample of 1000 individuals, what is the exact distribution of the number of people who carry this gene?

Answer: Binomial distribution with $p = 1/200; \quad n = 1000.$

II. (5 points) Use an approximation to this distribution to calculate the approximate probability that least people 10 carry the gene.

Answer: Since n is large and p is small, use Poisson with $\lambda = np = 5$

$P(X \geq 10) = 1 - P(X \leq 9) = 1 - F(9;5) = 1 - .968 = .032$