1. **Multiple Choice**.

(a) The proportion of commuters in Vancouver who use public transit to get to work is 20%. Let X be the number in a group of 400 commuters from Vancouver who use public transit to commute to work. What distribution does X follow? Check all that apply.

- ✓ $X \sim Bin(400, 0.2)$
- ____ $X \sim Bin(80, 0.2)$
- ____ $X \sim Geom(0.2)$
- ✓ $X \overset{approx}{\sim} N(80, 64)$
- ____ $X \overset{approx}{\sim} N(80, \frac{64}{400})$
- ____ None of the above

$n = 400$, $p = 0.2$

$np = 400 \times 0.2 = 80$, $n(1-p) = 320 \geq 5$

Normal approx. to binom.

$E(X) = np = 80$, $Var(X) = np(1-p) = 64$

For the following questions, please circle the most appropriate response.

(b) An ice cream stand reports that 12% of the cones they sell are new "jumbo" size cones. You want to see what a "jumbo" cone looks like, so you stand and watch the sales for a while. What is the probability that the first jumbo cone is the fourth cone you see them sell?

I. 8%
II. 33%
III. 40%
IV. 60%
V. 93%

X = 1st jumbo cone $\sim geo(p = 0.12)$

$P(X=4) = (1-0.12)^3 (0.12) = 0.08$

(c) Which of the following statements is true?

I. The true population standard deviation is an unbiased estimator for the sample standard deviation
II. The sampling distribution refers to the distribution of the values of items actually selected in a given sample
III. The sampling distribution of the mean is always normally distributed
IV. If the sample mean is an unbiased estimator for the population mean that implies the average sample mean, over all possible samples, equals the population mean
V. None of the above

(c) ~~(d)~~ The Central Limit Theorem states:

I. If n is large then the distribution of the sample will be approximately a normal distribution
II. The sampling distribution looks more like the population distribution as the sample size increases
III. If n is large, then the distribution of any statistic is approximately normal
IV. None of the above

Short Answer. Please show all steps and calculations. Be sure to define notation used and check assumptions that you use in your solutions unless otherwise specified.
Please specify the model and parameters used where necessary.

2. A researcher collected data on numbers of students entering the Woodward Library during various periods over a week's time. Between 12:00 and 12:10 pm (i.e. a 10 minute interval), an average of 40 students entered the library. The number of students entering the Woodward library can be modelled with a Poisson distribution.

(a) What is the probability that at least 1 student arrives at Woodward Library between 12:00 pm and 12:01 pm one day next week?

$X = \#$ of students entering library between 12:00 and 12:01 $\sim Pois\left(\frac{40}{10}\right)$

$P(X \geq 1) = 1 - P(X < 1)$

$= 1 - P(X = 0) = 1 - \frac{e^{-4}(4)^0}{0!} = 0.9817.$

(b) Suppose researchers begin observation at the Woodward Library entrance at exactly 12:00 pm. What is the probability of waiting more than 10 seconds for the first student to pass through the door?

$T =$ waiting time (minutes) until 1st student passes through $\sim exp(\lambda = 4)$

(10 seconds $\Rightarrow \frac{1}{6}$ min.)

$P(T > \frac{1}{6}) = 1 - F_T(\frac{1}{6})$

$F_T(t) = \int_0^t \lambda e^{-\lambda x}\, dx = 1 - e^{-\lambda t}$

$P(T > \frac{1}{6}) = 1 - (1 - e^{-4/6})$

$= 0.5134$

(c) What is the approximate probability that at least 30 students arrive between 12:00 and 12:10 pm one day next week?

$Y = \#$ students arriving between 12:00 and 12:10 $\sim Pois(40)$

$\lambda t = 40 \geq 20$ $\quad Y \overset{approx.}{\sim} N(\mu = 40, \sigma^2 = 40)$ (Normal approx. to Poisson)

$E(Y) = \lambda t$ $\quad Var(Y) = \lambda t$

$P(Y \geq 30) = P\left(Z \geq \frac{29.5 - 40}{\sqrt{40}}\right)$ continuity correction

$= 1 - P(Z < -1.66) = 1 - (1 - P(Z < 1.66)) = 0.9515$

Part b)

Alternate solution

W = # Students arriving in 10 seconds

$\sim \text{Pois}\left(\frac{4}{6}\right)$

$$P(W=0) = \frac{e^{-\frac{4}{6}}\left(\frac{4}{6}\right)^0}{0!} = 0.5134$$

3. A tall cup at Starbucks is designed to hold 355 mL of coffee. Suppose in reality, the amount of coffee poured into tall cups at Starbucks follow a normal distribution with mean 354 mL and standard deviation 1 mL. Assume the amount of coffee poured into tall cups is independent.

(a) Suppose you buy 3 randomly selected tall cups of coffee, what is the probability that exactly one of the 3 coffee cups is underfilled?

X = amount of coffee poured into tall cups $\sim N(354, 1^2)$

$P(\text{underfilled}) = P(X < 355) = P\left(Z < \frac{355-354}{1}\right) = P(Z < 1)$

≈ 0.84

(by 68-95-99.7 rule)

(Can also use $N(0,1)$ table)

Y = # of underfilled cups in 3

$Y \sim Bin(n=3, p=0.84)$

$P(Y=1) = \binom{3}{1} 0.84^1 (1-0.84)^{3-1} = 0.0645$

(b) Suppose you buy 40 randomly selected tall cups of coffee, what is the approximate probability that at most 26 are underfilled?

W = # underfilled cups in 40 $\sim Bin(n=40, p=0.84)$

$np = 33.6 \geq 5, \quad n(1-p) = 6.4 \geq 5$

$W \overset{\text{approx}}{\sim} N(\mu = 33.6, \sigma^2 = np(1-p) = 5.376)$

$P(W \leq 26) = P\left(Z \leq \frac{26.5 - 33.6}{\sqrt{5.376}}\right)$ continuity correction

$= P(Z \leq -3.06)$

$= 1 - 0.9989 = 0.0011$

4. The thickness of a protective coating applied to a conductor designed to work in corrosive conditions follows a uniform distribution between 10 and 16 microns.

(a) What are the mean and the standard deviation of the thickness of the protective coating?

X = thickness of protective coating $\sim U(10, 16)$

$$E(X) = \frac{10+16}{2} = 13 \qquad Var(X) = \frac{(16-10)^2}{12} = 3$$

$$SD(X) = \sqrt{3} = 1.732$$

(b) Suppose you randomly select 48 conductors. What is the approximate probability that in the 48 conductors, the mean thickness of the protective coating is greater than 13.5 minutes?

$X_i \sim U(10, 16) \qquad i = 1, \dots, 48$

$$\bar{X} = \frac{1}{48}\sum_{i=1}^{48} X_i \qquad E(\bar{X}) = 13$$

$$Var(\bar{X}) = \frac{3}{48} = \frac{1}{16}$$

$\bar{X} \sim N(13, \frac{1}{16})$ approx. by CLT $(n = 48 \geq 20)$

$$P(\bar{X} > 13.5) = P\left(Z > \frac{13.5-13}{\sqrt{1/16}}\right) = 1 - P(Z \leq 2)$$

$= 1 - 0.975$ (by approx. 68-95-99.7 rule)

$= 0.025$

5. Suppose the daily water consumption of a certain city follows a normal distribution with mean 30,000 m^3 and standard deviation 4000 m^3. The daily water supply follows a normal distribution with mean 40,000 m^3 and standard deviation 3000 m^3. Assume the consumption and supply of water are independent.

(a) What is the probability that on a randomly chosen day, the water consumption is more than 40,000 m^3?

Let X = daily water consumption $\sim N(30000, 4000^2)$

$$P(X>40000) = P\left(Z > \frac{40000-30000}{4000}\right) = P(Z>2.5)$$

$$= 1 - P(Z \leq 2.5) = 0.006$$

(b) If the conditions between days are independent, what is the probability the total water consumption in a month (assume 30 days) is more than 922,000 m^3?

Let X_i be the water concumption on the ith day

$T = X_1 + \ldots + X_{30}$ total water consumption in a month

$$E(T) = E(X_1 + \ldots + X_{20}) = E(X_1) + \ldots + E(X_{30}) = 30\,(30000) = 90000$$

$$Var(T) = Var(X_1 + \ldots + X_{30}) = Var(X_1) + \ldots + Var(X_{30}) \quad \text{independent } X_i$$

$$= 30\,Var(X) = 30 \times 4000^2$$

$$P(T>922000) = P\left(Z > \frac{922000 - 900000}{\sqrt{30 \times 4000^2}}\right) = P(Z>1.004)$$

$$= 1 - P(Z \leq 1.004)$$

$$= 0.159$$

(c) What is the probability of a water shortage in any given day?

Let Y = daily water supply $\sim N(40000, 3000^2)$

$P(X > Y) = P(X - Y > 0) = P$

$E(X - Y) = E(X) - E(Y)$

$= 30000 - 40000$

$= -10000$

$Var(X - Y) = Var(X) + Var(Y)$ X, Y independent

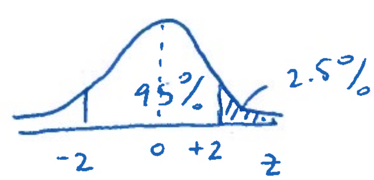
$= 4000^2 + 3000^2$

$= 25{,}000{,}000$

$P(X - Y > 0) = P\left(Z > \frac{0 + 10000}{5000}\right)$

$= P(Z > 2)$

≈ 0.025



by the 68-95-99.7 rule

or can look it up in the table.