A

5.4 46, 51, 55

5.5 58, 70, 73

38) Random variable $T_0 = X_1 + X_2$ can take the following values:

a) $0+0=0$; $0+1=1$; $0+2=2$; $1+2=3$; $2+2=4$

The probabilities are calculated using independence.

X_1	0	1	2
$p(x_1)$	0.2	0.5	0.3

$\mu = 1.1$ $\quad \sigma^2 = 0.49$

$P(T_0=0) = P(X_1=0, X_2=0) = P(X_1=0) \cdot P(X_2=0) = 0.2 \cdot 0.2 = 0.04$

$P(T_0=1) = P(X_1=1, X_2=0) + P(X_1=0, X_2=1)$

$= P(X_1=1)\,P(X_2=0) + P(X_1=0)\,P(X_2=1) = 0.5 \cdot 0.2 + 0.2 \cdot 0.5 = 0.2$

$P(T_0=2) = P(X_1=2, X_2=0) + P(X_1=0, X_2=2) + P(X_1=1, X_2=1)$

$= P(X_1=2)\,P(X_2=0) + P(X_1=0)\,P(X_2=2) + P(X_1=1)\,P(X_2=1)$

$= 0.3 \cdot 0.2 + 0.2 \cdot 0.3 + 0.5 \cdot 0.5 = 0.37$

$P(T_0=3) = P(X_1=1, X_2=2) + P(X_1=2, X_2=1) = P(X_1=1)\,P(X_2=2) + P(X_1=2)\,P(X_2=1)$

$= 0.5 \cdot 0.3 + 0.3 \cdot 0.5 = 0.3$

$P(T_0=4) = P(X_1=2, X_2=2) = P(X_1=2) \cdot P(X_2=2) = 0.3 \cdot 0.3 = 0.09$

t_0	0	1	2	3	4
$p(t_0)$	0.04	0.2	0.37	0.3	0.09

b) $E(T_0) = \mu_{T_0} = 0 \cdot 0.04 + 1 \cdot 0.2 + 2 \cdot 0.37 + 3 \cdot 0.3 + 4 \cdot 0.09 = 2.2$

Expectations of random variables X_1 and X_2 are the same:

$E(X_1) = E(X_2) = \mu = 0 \cdot 0.2 + 1 \cdot 0.5 + 2 \cdot 0.3 = 1.1$

$\mu_{T_0} = 2.2$ which is 2 times the population mean μ.

c) Expectation of random variable T_0^2: $E(T_0^2) = 0^2 \cdot 0.04 + 1^2 \cdot 0.2 + 2^2 \cdot 0.37 + 3^2 \cdot 0.3$
$+ 4^2 \cdot 0.09 = 5.82$

Variance of random variable T_0 is $\sigma^2_{T_0} = E(T_0^2) - [E(T_0)]^2 = 5.82 - 2.2^2 = 0.98$

Expectation of " " X_1^2 is $E(X_1^2) = 0^2 \cdot 0.2 + 1^2 \cdot 0.5 + 2^2 \cdot 0.3 = 1.7$

Variance of population ~~random~~ " X_1 is $\sigma^2 = \sigma^2_{X_1} = E(X_1^2) - [E(X_1)]^2 = 1.7 - 1.1^2 = 0.49$

So, " " " " $\sigma^2_{T_0}$ is 2 times than the variance of population ~~random var~~. σ^2

d) Random variable $T_0 = X_1 + X_2 + X_3 + X_4$

$E(X_1) = E(X_2) = E(X_3) = E(X_4) = 1.1$

$E(T_0) = E(X_1 + X_2 + X_3 + X_4) = E(X_1) + E(X_2) + E(X_3) + E(X_4) = 4 \cdot 1.1 = 4.4$

$V(T_0) = \underbrace{V(X_1 + X_2 + X_3 + X_4) = V(X_1) + V(X_2) + V(X_3) + V(X_4)}_{\text{iff they are independent, you can do this}} = 4 \cdot 0.49 = 1.96$

c) The only way that the total number of lights at which a stop is required is 8, is if at each way (random variable) the required stops are 2.

$P(T_o = 8) = P(X_1 = 2, X_2 = 2, X_3 = 2, X_4 = 2) = P(X_1 = 2) . P(X_2 = 2) , P(X_3 = 2) . P(X_4 = 2)$

$= 0.3 . 0.3 . 0.3 . 0.3 = 0.0081$

$P(T_o = 7) = \cancel{P(X=} (0.3 . 0.3 . 0.3 . 0.5) + (0.3 . 0.3 . 0.5 . 0.3) + (0.3 . 0.5 . 0.3 . 0.3)$

$+ (0.5 . 0.3 . 0.3 . 0.3) = 4 (0.3^3 . 0.5) = 0.054$

$P(T_o \geq 7) = P(T_o = 7) + P(T_o = 8) = 0.054 + 0.0081 = 0.0621$

41)

a)

x	1	2	3	4
$p(x)$	0.4	0.3	0.2	0.1

First data value	Second data value	Sample mean	Probability	Sample range
1	1	1	0.16	0
1	2	1.5	0.12	1
1	3	2	0.08	2
1	4	2.5	0.04	3
2	1	1.5	0.12	1
2	2	2	0.09	0
2	3	2.5	0.06	1
2	4	3	0.03	2
3	1	2	0.08	2
3	2	2.5	0.06	1
3	3	3	0.04	0
3	4	3.5	0.02	1
4	1	2.5	0.04	3
4	2	3	0.03	2
4	3	3.5	0.02	1
4	4	4	0.01	0

Above table contain every random sample with 2 data values $n=2$ from the set $\{1,2,3,4\}$ (selection of the same value is allowed)

Sample mean is sum of all values divided by number of values.

Probability is the product of the possibilities associated with 2 data values.

~~Sample range~~ is the difference between the largest and smallest value.

Sample mean	Probability
1	0.16
1.5	0.12 + 0.12 = 0.24
2	0.08 + 0.09 + 0.08 = 0.25
2.5	0.2
3	0.1
3.5	0.04
4	0.01

Probability of the sample mean is the sum of the probabilities in the previous table that lead to the same sample mean $\bar{x}$

b) For disjoint/mutually exclusive events: $P(\bar{x} \leq 2.5) = P(x=1) + P(x=1.5) + P(x=2) + P(x=2.5) = 0.16 + 0.24 + 0.25 + 0.2 = 0.85$

c)

Sample range	Probability
0	0.3 = 0.16 + 0.09 + 0.04 + 0.02
1	0.4
2	0.22
3	0.08

Probability of sample range is the sum of the probabilities in the first table that lead to the same sample mean.

d) Sample of size $n=4$ with sample mean of atmost 1.5, need to have the properties that the sum of all data values is atmost 6 (because the sample mean is the sum of all data values divided by $n=4$): {(1,1,1,1), (1,1,1,2), (1,1,2,1), (1,2,1,1), (2,1,1,1), (1,1,2,2), (1,2,1,2), (2,1,1,2), (1,2,2,1), (2,1,2,1), (2,2,1,1), (1,1,1,3), (1,3,1,1), (3,1,1,1)}

~~Possibilities~~ Probability: product of probability of each of the 4 values:
{0.0256, 0.0192, 0.0192, 0.0192, 0.0192, 0.0144, 0.0144, 0.0144, 0.0144, 0.0144, 0.0144, 0.0128, 0.0128, 0.0128, 0.0128}

$P(\bar{x} \leq 1.5) = 0.0256 + 4(0.0192) + 6(0.0144) + 4(0.0128) = 0.24$

46) The sampling distribution of $\bar{X}$ is centered at $E(\bar{X}) = \mu = 12$ cm, and the

a) standard deviation of the $\bar{X}$ distribution is $\sigma_{\bar{X}} = \frac{\sigma_{\bar{x}}}{\sqrt{n}} = \frac{0.04}{\sqrt{16}} = 0.01$ cm

b) With $n=64$, the sample distribution of $\bar{X}$ is still centered at $E(\bar{X}) = \mu = 12$ cm, but the standard deviation of the $\bar{X}$ distribution is $\sigma_{\bar{x}} = \frac{\sigma_{\bar{x}}}{\sqrt{n}} = \frac{0.04}{\sqrt{64}} = 0.005$ cm

c) $\bar{X}$ is more likely to be within 0.01 cm of the mean (12 cm) with the second, larger, sample. This is due to the decreased variability of $\bar{X}$

that comes with a larger sample size.

51) $\mu = 10$ minutes, $\sigma = 2$ minutes, expected value of random variable $\bar{X}$ (sample average) is $\mu_{\bar{x}} = 10$

$\sigma_{\bar{x}} = \frac{1}{\sqrt{n}} \cdot \sigma = \frac{1}{\sqrt{5}} \cdot 2 = 0.894$, when $n = 5$

$\sigma_{\bar{x}} = \frac{1}{\sqrt{n}} \cdot \sigma = \frac{1}{\sqrt{6}} \cdot 2 = 0.816$ " $n = 6$

From appendix table A3

when $n=5$, $P(\bar{X} \leq 11) = P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \leq \frac{11-10}{0.894}\right) = P(Z \leq 1.12) = 0.8686$

when $n=6$, $P(\bar{X} \leq 11) = P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \leq \frac{11-10}{0.816}\right) = P(Z \leq 1.22) = 0.8888$

Assume the results of the 2 days are independent (which seems reasonable), the probability of the sample average at most 11 min on both days is $P(\bar{X} \leq 11) = P_1(\bar{X} \leq 11) \cdot P_2(\bar{X} \leq 11) = 0.8686 \cdot 0.8888 = 0.772$

55) random variable X with Poisson Distribution with $\mu > 0$, $E(X) = V(X) = \mu$

a) The distribution, Poisson, has mean and variance equal to 50. For the probability that the random variable X to be between 2 given values:

$P(35 < X < 70) \approx P\left(\frac{35 - 0.5 - E(X)}{\sigma_X} < \frac{X - E(X)}{\sigma_X} < \frac{70 + 0.5 - E(X)}{\sigma_X}\right)$

correction (continuity) for Poisson distribution (-0.5).

$= P\left(\frac{34.5 - 50}{\sqrt{50}} < Z < \frac{70.5 - 50}{\sqrt{50}}\right) = P(-2.19 < Z \leq 2.9)$

$= P(Z < 2.9) - P(Z < -2.19) = \phi(2.9) - \phi(-2.19)$.

$= 0.9981 - 0.0143 = 0.9838$ → Appendix Table A.3

Z has a standard normal distribution
- mean 0, standard dev 1 (using CLT)

b) In the 5 day period, the mean changes 5 times: $\mu_5 = 50.5 = 250$, which is also variance of such random variable.

Like in a, $P(225 < X \; 275) \approx P\left(\frac{225 - 0.5 - E(X)}{\sigma_X} < \frac{X - E(X)}{\sigma_X} < \frac{275 + 0.5 - E(X)}{\sigma_X}\right)$

$= P\left(\frac{224.5 - E(X)}{\sigma_X} < Z < \frac{275.5 - E(X)}{\sigma_X}\right) = P(-1.61 < Z \; 1.61)$

$= P(Z < -1.61) - P(Z < 1.61) = \phi(1.61) - \phi(-1.61) = 0.9463 - 0.0537 = 0.8926$

58) Volume $= 27X_1 + 125X_2 + 512X_3$

$E(V_0) = E(27X_1 + 125X_2 + 512X_3) = 27E(X_1) + 125E(X_2) + 512E(X_3)$

$= 27 \cdot 200 + 125 \cdot 250 + 512 \cdot 100 = 87850$

Variance of Volume $V(V_0) = V(27X_1 + 125X_2 + 512X_3) = 27^2 V(X_1) + 125^2 V(X_2)$

$+ 512^2 V(X_3) = 27^2 \cdot 10^2 + 125^2 \cdot 12^2 + 512^2 \cdot 8^2 = 19,100,116$

Because given random variables are independent.

b) The expected value would still be the same no matter the independence. Variance would change because the covariances now also contribute to the variance.

70) The expected values of Bernoulli random variable is 0.5. This is because
a) such random variable is either $p(0) = 0.5$ or $p(1) = 0.5$.

$E(Y_i) = 0 \cdot 0.5 + 1 \cdot 0.5 = 0.5$

Expected value of W $E(W) = E(1 \cdot Y_1 + 2Y_2 + \cdots + nY_n) = 1 \cdot E(Y_1) + 2E(Y_2) + \cdots + n \cdot E(Y_n) = \sum_{k=1}^{n} k \cdot E(Y_k) = \sum_{k=1}^{n} k \cdot 0.5 = 0.5 \sum_{k=1}^{n} k = \frac{n(n+1)}{4}$

b) Variance of Bernoulli random variable $V(Y_i) = E(Y_i^2) - [E(Y_i)]^2$

$= (0^2 \cdot 0.5 + 1^2 \cdot 0.5) - 0.5^2 = 0.25$

Given random variables Y_i are independent, : $V(W) = V(1 \cdot Y_1 + 2 \cdot Y_2 + \cdots + n \cdot Y_n)$

$= 1^2 \cdot V(Y_1) + 2^2 \cdot V(Y_2) + \cdots + n^2 V(Y_n) = \sum_{k=1}^{n} k^2 V(Y_k) = \sum_{k=1}^{n} k \cdot 0.25 = \frac{1}{4} \sum_{k=1}^{n} k^2$

$= \frac{n(n+1)(2n+1)}{24}$

73) $\mu_x = 105$, $\sigma_x = 8$, $\mu_y = 100$, $\sigma_y = 6$, $n_x = 40$, $n_y = 35$.

Central limit theorem: If the sample size is large (≥ 30), then the sampling distribution of the sample means $\bar{x}$ and $\bar{y}$ are approximately normal.

Since sample sizes $n_x = 40$ and $n_y = 35$ (≥ 30), we know that sampling distribution of the sample mean $\bar{x}$ is approximately normal, by CLT.

The sampling distribution of the sample mean $\bar{x}$ and $\bar{y}$ has mean μ and standard dev $\frac{\sigma}{\sqrt{n}}$:

$\mu_{\bar{x}} = \mu_x = 105$; $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{8}{\sqrt{40}} = \frac{2\sqrt{10}}{5} \approx 1.2649$

$\mu_{\bar{y}} = \mu_y = 100$; $\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} = \frac{6}{\sqrt{35}} \approx 1.0142$

b) For linear combination $W = aX_1 + bX_2$, the following properties hold for mean, variance and standard deviation:

$\mu_W = a\mu_1 + b\mu_2$

$\left.\begin{aligned} \sigma_W^2 &= a^2\sigma_1^2 + b^2\sigma_2^2 \\ \sigma_W &= \sqrt{a^2 \cdot \sigma_1^2 + b^2 \cdot \sigma_2^2} \end{aligned}\right\}$ (If X_1 and X_2 are independent)

Assume that $\bar{X}$ and $\bar{Y}$ are independent:

$\mu_{\bar{x}-\bar{y}} = \mu_{\bar{x}} - \mu_{\bar{y}} = 105 - 100 = 5$

$\sigma_{\bar{x}-\bar{y}} = \sqrt{\sigma_{\bar{x}}^2 + (-1)^2\sigma_{\bar{y}}^2} = \sqrt{1.2649^2 + 1.0142^2} \approx 1.6213$

$\bar{X} - \bar{Y}$ is approximately normally distributed, because $\bar{X}$ and $\bar{Y}$ are approximately normally distributed (assuming they are independent).

c) Standardized score is the value x decreased by the mean and divided by standard deviation.

$$P(-1 \le \bar{X} - \bar{Y} \le 1) = P(-3.7 < Z < -2.47) = P(Z < -2.47) - P(Z < -3.7)$$

$$= 0.0068 - 0 = 0.0068 = 0{,}68\%$$ → Appendix Table A.3

$$Z = \frac{x - \mu}{\sigma} = \frac{1 - 5}{1.6213} = -2.47$$

$$Z = \frac{x - \mu}{\sigma} = \frac{-1 - 5}{1.6213} = -3.7$$

d) $$Z = \frac{x - \mu}{\sigma} = \frac{10 - 5}{1.6213} \approx 3.08$$

$$P(\bar{X} - \bar{Y} \ge 10) = P(Z > 3.08) = 1 - P(Z < 3.08) = 1 - 0.9990 = 0.001 = 0{,}1\%$$

Since this probability is very small, we would doubt that $\mu_1 - \mu_2 = 5$ if we obtained that $\bar{X} - \bar{Y} \ge 10$.