(1) $\Omega = \{a, b, c\}$, $p(a) = 0'1$, $p(b) = 0'2$ and $p(c) = 0'7$.

(a) $E[f(x)] = \sum_{a,b,c} f(x)\,p(x)$

$= f(a).p(a) + f(b).p(b) + f(c).p(c)$

$= 10.0'1 + 5.0'2 + \frac{10}{7}.0'7$

$= 1 + 1 + 1$

$= 3$

$\therefore E[f(x)] = 3$.

(b) $E[1/p(x)] = \sum_{a,b,c} f(x).p(x)$ where $f(x) = \frac{1}{p(x)}$

$= \cancel{\sum_{a,b,c} \frac{1}{p(x)}\, p(x)} = \sum_{a,b,c} \left(\frac{1}{p(x)}\right) \cdot p(x)$

$= \frac{1}{p(a)} \cdot p(a) + \frac{1}{p(b)} \cdot p(b) + \frac{1}{p(c)} \cdot p(c)$

$= 1 + 1 + 1$

$= 3$.

(c) $E[1/p(x)]$ for arbitrary pmf p,

$\sum f(x).p(x)$ for any a, b, c

or, $\sum \frac{1}{p(x)} \cdot p(x) = \sum 1$ or 1 for any a, b, c.

(d) $E[f(x)^2] = \sum_{a,b,c} f(x)^2 p(x)$

$$= f(a)^2 p(a) + f(b)^2 p(b) + f(c)^2 \cdot p(c)$$

$$= (10)^2 \cdot 0.1 + 5^2 \cdot 0.2 + \left(\frac{10}{7}\right)^2 \cdot 0.7$$

$$= 10 + 5 + \frac{10}{7}$$

$$\approx 16.43$$

Thus, $E[f(x)^2] = 16.43$

Now, $E[f(x)]^2 = \sum_{a,b,c} (f(x) \cdot p(x))^2$

$$= [f(a) \cdot p(a) + f(b) \cdot p(b) + f(c) \cdot p(c)]^2$$

$$= [(10) \cdot (0.1) + 5 \cdot (0.2) + \left(\frac{10}{7}\right) \cdot (0.7)]^2$$

$$= \cancel{100 \cdot 0.01 + 25 \cdot 0.04 + \frac{100}{49} \cdot \frac{49}{100}}$$

$$= [1 + 1 + 1]^2$$

$$= 3^2 = 9$$

And, $E[f(x)]^2 = 9$

2. (a) $p(A) = 0.75$, $p(B) = 0.5$, $p(C) = 0.25$.

All 3 coins are flipped once.

We know, expected values,

$$E[x] = \sum_{\substack{\cancel{x \in 0,1,2,3} \\ x \in 1}} x \cdot p(x)$$

$$= 1 \cdot p(A) + 1 \cdot p(B) + 1 \cdot p(C)$$

$$= 1 \cdot 0.75 + 1 \cdot 0.5 + 1 \cdot 0.25$$

$$= 0.75 + 0.5 + 0.25$$

$$= 1.5$$

Ans: Expected value of x being heads is 1.5.

(b) Probability of choosing each coin,
$p = 1/3$.

Probability of choosing each coin, when 3 out of 5 flips are heads:

$$p(D \mid A) = {}^5C_3 \cdot (0.75)^3 \cdot (1-0.75)^2$$
$$= 10 \cdot (0.75)^3 (0.25)^2$$
$$= 0.2604$$

$$p(D \mid B) = {}^5C_3 \, (0.5)^3 (1-0.5)^2$$
$$= 10 \cdot (0.5)^5$$
$$= 0.3125$$

$$p(D \mid C) = {}^5C_3 \, (0.25)^3 (1-0.25)^2$$
$$= 10 \cdot (0.25)^3 (0.75)^2$$
$$= 0.088$$

Probability of choosing coin C, given D = 3 Heads and 2 tails.

$$p(C \mid D) = \frac{p(D \mid C) \cdot p(C)}{p(D)}$$
$$= \frac{0.088 \cdot \cancel{1/3}}{\cancel{1/3}\,[0.264 + 0.3125 + 0.088]}$$
$$= 0.132$$

Ans: probability of choosing C = 0.132.

3. Given, cost of an electrical breakdown of duration x, $f(x) = x^3$.

Following the uniform distribution

$$p(x) = \begin{cases} \frac{1}{10} & \text{if } 0 < x < 10 \\ 0 & \text{otherwise} \end{cases}$$

we get, expected cost of an electrical breakdown

$$E[f(x)] = \int f(x)\, p(x)\, dx$$

$$= \int_0^{10} x^3 \cdot \frac{1}{10} \cdot dx + \int_{10}^{\infty} x^3 \cdot 0 \cdot dx$$

$$= \frac{1}{10} \int_0^{10} x^3\, dx + 0$$

$$= \frac{1}{10} \left[\frac{x^4}{4} \right]_0^{10} = \frac{1}{40} (10)^4$$

$$= \frac{\cancel{1000}^{\,250}}{\cancel{4}} = 250$$

Ans: Expected cost of an ~~ete~~ electrical breakdown is 250.

4 (b) Variance, $\sigma^2 = 10$ and $\mu = 0$. for 10 samples.

We got the means from code :

$$[0.092, 0.136, 0.553, 0.3527, -0.101]$$

Variance, $\bar{V} = \frac{1}{n-1} \sum_{i=1}^{n} (M_i - \bar{M})^2$

Here, sample average of the means, $\bar{M} = 0.2065$.

$\bar{V} = \frac{1}{4} [(0.092 - 0.2065)^2 + (0.136 - 0.2065)^2 +$
$(0.553 - 0.2065)^2 + (0.3527 - 0.2065)^2 + (-0.101 - 0.2065)^2]$

$= \frac{1}{4} [\cancel{0.0131 + 0.0497 + 0.12 + 0.095} \; {}^{+0.0214}]$
$[0.254]$

$= 0.064$

<u>Ans</u>: Unbiased sample variance $\bar{V} = 0.064$.

(c) Means from the 100 samples ~~as~~ from code:

$$[-0.059, 0.1058, 0.0677, -0.1398, 0.0817]$$

Sample Variance, $\bar{V} = \frac{1}{\cancel{99.4}\; n-1} [(M_i - \bar{M})]^2$

Here, $\bar{M} = 0.0113$.

Variance, $\bar{V} = \frac{1}{5-1} [0.045] = \frac{1}{4}(0.045)$ [Similar to 4b]

$= \cancel{0.112}$
$= 0.0112$

Here, we can see that increasing the ~~sa~~ number of samples decreases the variance.

Because, Variance for 10 samples, 0.064 > 0.0112 variance for a 100 samples.

This is because as we get more samples, the closer $\bar{x}$ is randomly sampled closer to the mean, which reduces variance.

(d) $\sigma^2 = 10.0$ [Variance]

Sample average, $M =$ ~~0.1138~~ 0.469

for 95% CI, $\delta = 0.05$

This is a Gaussian distribution, so, the true mean

$$\mu = \left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$$

Here, $\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{10}}{\sqrt{30}} = 0.577.$

$$\therefore \mu \in [0.469 - 0.577, 0.469 + 0.577]$$

$$= [-0.108, 1.046]$$

Therefore, the 95% confidence interval around $M = 0.469$ is $(-0.108, 1.046)$.

(e) $\sigma^2 = 10$, and not a Gaussian distribution.

for 95% confidence interval, $\delta = 0.05$,

$$\epsilon = \sqrt{\frac{\sigma^2}{n\delta}} = \sqrt{\frac{\cancel{0.05^2}\ 10^2}{30 \cdot 0.05}} = \frac{10}{\sqrt{1.5}}$$

or, $\epsilon = 8.165$.

Our mean (sample average) M from (d) $= 0.469$.

$\therefore$ True mean $\mu \in [0.469 - 8.165, 0.469 + 8.165]$

or, $\mu \in [-7.7, 8.634]$

Thus, 95% confidence interval around $M = 0.469$ for non-Gaussian data is $[-7.7, 8.634]$

(5) (a) Unbiased Variance ~~Estimator~~, $\bar{V} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Biased variance, $\bar{V}_b = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Now, expectation of biased ~~vara~~ variance estimator, $E[\bar{V}_b] = E\left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2\right]$

using property $E[cX] = cE[X]$ or, $E[\bar{V}_b] = \frac{1}{n} E\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right]$

also, unbiased expectation, $E[\bar{V}] = \sigma^2$

or, $E\left[\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2\right] = \sigma^2$

or, $\frac{1}{n-1} E\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right] = \sigma^2$

or, $E\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right] = (n-1)\sigma^2 \quad \ldots (1)$

Now, again from $E[\bar{V}_b] = \frac{1}{n} E\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right]$

$= \frac{1}{n}(n-1)\sigma^2$ [from (1)]

$= \left(\frac{n}{n} - \frac{1}{n}\right)\sigma^2$

$\therefore E[\bar{V}_b] = \left(1 - \frac{1}{n}\right)\sigma^2$

<u>Ans:</u> $E[\bar{V}_b] = \left(1 - \frac{1}{n}\right)\sigma^2$

[Shown]

(b) $Var[\bar{x}] = \frac{1}{n}\sigma^2$

Given, ~~Cheby~~ $Var[\bar{v}] = \frac{2(n-1)}{n^2}\sigma^4$

We know, Chebyshev's inequality:

$$Pr(|x-\mu| \geq s) \leq \frac{\sigma^2}{n s^2}$$

Also, $Pr(|V - \overset{E[V]}{\cancel{E[\,]}}| \geq s) \leq \frac{V}{s^2}$ as stated in question for random variable x, for any $s>0$

or $Pr(|V - E[\bar{v}]| \geq s) \leq \frac{\sigma^2}{s^2}$ [variance $= \sigma^2$]

taking $s = n\sigma^2$,

$$Pr(|V - E[\bar{v}]| \geq n\sigma^2) \leq \frac{\sigma^2}{n^2\sigma^4}$$

or $Pr(|V - E[\bar{v}]| \geq n\, Var[\bar{V}]) \leq \frac{1}{n^2\sigma^2}$

or, $Pr(|V - E[\bar{v}]| \geq n \cdot \frac{2(n-1)}{n^2}\sigma^4) \leq \frac{1}{n^2\sigma^2}$

or, $Pr(|V - E[\bar{v}]| \geq 2(1-\frac{1}{n})\sigma^4) \leq \frac{1}{n^2\sigma^2}$

here, as $\lim n\to\infty$, we get

$\lim_{n\to\infty} \frac{1}{n^2\sigma^2} = 0$, or close to 0 as n becomes larger.

Also, $\lim_{n\to\infty} 2(1-\frac{1}{n})\sigma^4 = 2\sigma^4$, or close to $2\sigma^4$ as n becomes larger. Thus, for $\epsilon = 2(1-\frac{1}{n})\sigma^4$, our confidence interval is tighter around the sample ~~variance~~ ~~$\bar{V}$~~ mean $\bar{x}$, as ϵ is larger for the sample variance $\bar{v}$.