Student Information

Name : Mithat Can Timurcan

ID : 2581064

Answer 1

a)

- To find the value of k, we need to make sure that the total probability is equal to 1.

$$\mathbf{P}\{(X,Y) \in A\} = \iint\limits_{(x,y)\in A} f_{X,Y}(x,y) \, dx \, dy = 1 \text{ where } A = [0,1] \times [0,1]$$

- Which can be written as:

$$\iint\limits_{(x,y)\in A} f_{X,Y}(x,y) \, dx \, dy = \int_0^1 \int_0^1 (x + ky^3) \, dx \, dy = \int_0^1 \left(\frac{x^2}{2} + ky^3 x\right)\Big|_0^1 \, dy$$

$$= \int_0^1 \left(\frac{1}{2} + ky^3\right) \, dy = \left(\frac{y}{2} + \frac{ky^4}{4}\right)\Big|_0^1 = \left(\frac{1}{2} + \frac{k}{4}\right) = 1$$

- Solving for k we get:

$$\left(\frac{1}{2} + \frac{k}{4}\right) = 1 \rightarrow k = 2$$

b)

- To find the probability of $\mathbf{P}\{X = \frac{1}{2}\}$, we need to integrate over the joint density function in the following way:

$$\mathbf{P}\{(X,Y) \in A\} = \iint\limits_{(x,y)\in A} f_{X,Y}(x,y) \, dx \, dy \text{ where } A = \{\frac{1}{2}\} \times [0,1]$$

$$\mathbf{P}\{X = \frac{1}{2}\} = \mathbf{P}\{X = \frac{1}{2}, 0 \le Y \le 1\} = \int_0^1 \int_{\frac{1}{2}}^{\frac{1}{2}} (x + 2y^3) \, dx \, dy = 0$$

c)

- To find the probability of $\mathbf{P}\{0 \le X \le \frac{1}{2}, 0 \le Y \le \frac{1}{2}\}$, we need to integrate over the joint density function in the following way:

$$\mathbf{P}\{(X,Y) \in A\} = \iint\limits_{(x,y)\in A} f_{X,Y}(x,y) \, dx \, dy \text{ where } A = [0, \frac{1}{2}] \times [0, \frac{1}{2}]$$

$$\iint\limits_{(x,y)\in A} f_{X,Y}(x,y)\, dx\, dy = \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} (x + 2y^3)\, dx\, dy = \int_0^{\frac{1}{2}} \left(\frac{x^2}{2} + 2xy^3 \right) \Big|_0^{\frac{1}{2}} dy$$

$$= \int_0^{\frac{1}{2}} \left(\frac{1}{8} + y^3 \right) dy = \left(\frac{y}{8} + \frac{y^4}{4} \right) \Big|_0^{\frac{1}{2}} = \left(\frac{1}{16} + \frac{1}{64} \right) = \frac{5}{64}$$

Answer 2

a)

- In order to find the marginal pdf of Y, we need to do the following:

$$f_Y(y) = \int_0^\infty f_{X,Y}(x,y)\, dx = \int_0^\infty \left(\frac{e^{(-y-\frac{x}{y})}}{y} \right) dx = \frac{e^{-y}}{y} \int_0^\infty e^{-\frac{x}{y}}\, dx$$

$$\frac{e^{-y}}{y} \left(-ye^{-\frac{x}{y}} \Big|_0^\infty \right) = \frac{e^{-y}}{y}(0 - (-y)) = \frac{e^{-y}}{y} \cdot y = e^{-y}$$

- We can see that $f_Y(y)$ is in the form of $\lambda e^{\lambda y}$ with $\lambda = 1$. Therefore, we can say that our variable Y follows the exponential distribution.

b)

- We can find the expected value of Y with the following formula:

$$\mathbf{E}(Y) = \mu_Y = \int_0^\infty y f_Y(y)\, dy = \int_0^\infty y e^{-y}\, dy$$

- We also know that if a variable follows the exponential distribution, it's expected value can be computed using the λ value:

$$\mathbf{E}(Y) = \mu_Y = \frac{1}{\lambda}$$

- Since we've found the λ value as 1, we can see that our expected value of Y, $\mathbf{E}(Y)$ is equal to 1.

Answer 3

a)

- We can use binomial distribution in this problem. Since we know that 10% of the soldiers belong to the naval forces, we can take the success value p as 0.1. We want to make sure

that 9% of the 1000 soldiers belong to the naval forces so there should be at least 90 soldiers. We can calculate this in the following way:

$$\mathbf{P}\{X \geq 90\} = 1 - \mathbf{P}\{X \leq 89\} = 1 - \sum_{k=0}^{89} \binom{1000}{k} p^k (1-p)^{1000-k} \approx 0.867$$

calculated using the binocdf function on octave.

- We also know that for large values of n and moderate values of p, Normal distribution tends to approximate Binomial distribution with the mean of $\mu = np$ and standard deviation of $\sigma = \sqrt{np(1-p)}$. However, we also need to apply continuity correction since we're approximating a discrete distribution by a continuous distribution so we're going to expand our interval by 0.5 units, making it 89.5. Then we can use the `normcdf(x, `μ`, `σ`)` on Octave.

$$\mathbf{P}\{X \geq 90\} = 1 - \mathbf{P}\{X \leq 89.5\}$$
$$1 - \texttt{normcdf(89.5, 1000*0.1, sqrt(1000*0.1*0.9))} \approx 0.866$$

b)

- We can use the same approach we used in part a). However, this time our pool consists of 2000 soldiers so we need to make sure that there are at least 180 that belong to the naval forces. We can achieve this in the following way:

$$\mathbf{P}\{X \geq 180\} = 1 - \mathbf{P}\{X \leq 179\} = 1 - \sum_{k=0}^{179} \binom{2000}{k} p^k (1-p)^{2000-k} \approx 0.939$$

calculated using the binocdf function on octave.

- We can also use Normal approximation to Binomial like we did in the previous part:

$$\mathbf{P}\{X \geq 180\} = 1 - \mathbf{P}\{X \leq 179.5\}$$
$$1 - \texttt{normcdf(179.5, 2000*0.1, sqrt(2000*0.1*0.9))} \approx 0.937$$

- According to the Central Limit Theorem, as the sample size increases, the distribution of the sample proportion approaches a normal distribution. Consequently, the variability in the proportion of naval soldiers across various samples diminishes with larger sample sizes. This reduced variability increases the likelihood of consistently meeting or surpassing the minimum percentage requirement.

Answer 4

a)

- First of all, we should standardize our variable, (let us denote with T) with a mean μ of 65 years and a standard deviation σ of 6 years:

$$\mathbf{P}\{60 < T < 75\} = \mathbf{P}\left\{\frac{60 - \mu}{\sigma} < \frac{T - \mu}{\sigma} < \frac{75 - \mu}{\sigma}\right\}$$

$$\mathbf{P}\left\{\frac{60 - 65}{6} < Z < \frac{75 - 65}{6}\right\} = \mathbf{P}\{-0.833 < Z < 1.667\}$$

$$\Phi(1.667) - \Phi(-0.833) \approx 0.750$$

calculated using the normcdf function on octave.

b)

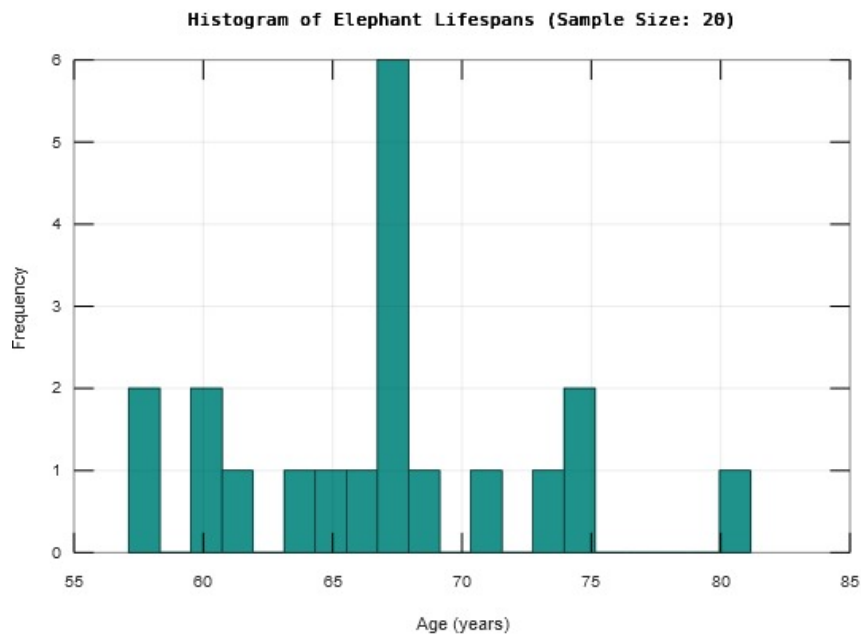Figure 1: Histogram when the sample size is 20.

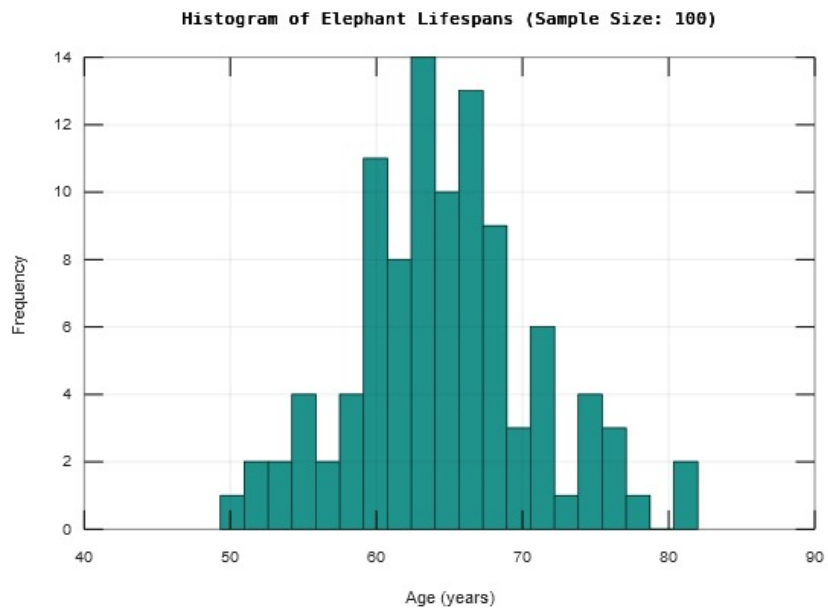Histogram of Elephant Lifespans (Sample Size: 1000)
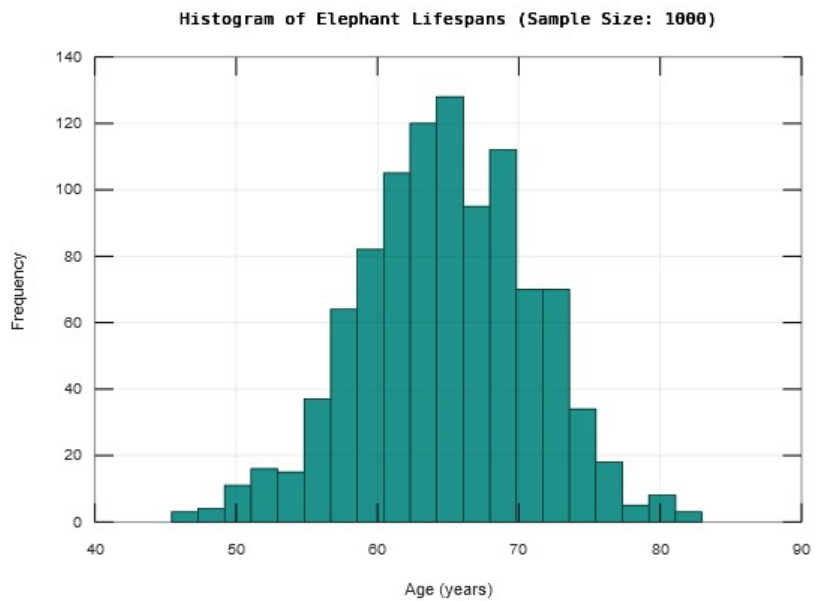
Figure 2: Histogram when the sample size is 100.

Figure 3: Histogram when the sample size is 1000.

- These histograms verify our knowledge that as the sample size increases, the distribution of the sample proportion approaches a normal distribution according to the Central Limit Theorem.

```
octave:14> mu = 65;
sigma = 6;
sample_sizes = [20, 100, 1000];

for i = 1:length(sample_sizes)

    sample = normrnd(mu, sigma, [sample_sizes(i), 1]);

    figure;
    hist(sample, 20);
    title(['Histogram of Elephant Lifespans (Sample Size: ' num2str(sample_sizes(i)) ')']);
    xlabel('Age (years)');
    ylabel('Frequency');
    grid on;
end
```

Figure 4: Octave/Matlab code for the plots.

c)

```
octave:7> mu = 65;
sigma = 6;
range_min = 60;
range_max = 75;
sample_size = 100;
iterations = 1000;

count_70_percent = 0;
count_85_percent = 0;


for iter = 1:iterations
    sample = normrnd(mu, sigma, [sample_size, 1]);

    count_in_range = sum(sample >= range_min & sample <= range_max);

    if count_in_range >= 0.7 * sample_size
        count_70_percent = count_70_percent + 1;
    end

    if count_in_range >= 0.85 * sample_size
        count_85_percent = count_85_percent + 1;
    end
end

disp(['Number of simulations where at least 70% of the elephants had a lifespan in the given range: ' num2str(count_70_percent)]);
disp(['Number of simulations where at least 85% of the elephants had a lifespan in the given range: ' num2str(count_85_percent)]);
Number of simulations where at least 70% of the elephants had a lifespan in the given range: 905
Number of simulations where at least 85% of the elephants had a lifespan in the given range: 8
```

Figure 5: Octave/Matlab code and the output for the simulations.

- In part a), we found that approximately 75% of the elephants will live more than 60 years, but less than 75 years. Looking to our simulations, in 905 of them, we observe at least 70% of elephants having a lifespan in the given range. This result aligns closely with our expected percentage of 75%. It suggests that our model's predictions are consistent with reality in this case. However, in only 8 of the simulations, we observe that at least 85% of the elephants having a lifespan in the given range. This result is significantly lower than our expected percentage of 75%. It indicates that such a high concentration of elephants within the specified age range is less likely to occur based on our model's predictions.