DAI Assignment 1

Tamanna Kumari, 24B1015

Sagar V, 24B1021

Videep Reddy Jalapally, 24B1037

Question 3

Bonferroni's inequality states that for two events A and B:

$$P(A \cap B) \geq P(A) + P(B) - 1$$

We know that $P(A) \geq 1 - q_1$ and $P(B) \geq 1 - q_2$. Thus, we can write:

$$P(A \cap B) \geq (1 - q_1) + (1 - q_2) - 1$$

$$P(A \cap B) \geq 1 - (q_1 + q_2)$$

Hence, Proven that $P(A \cap B) \geq 1 - (q_1 + q_2)$.

Question 4

Labelling events:

R Event that a bus in town is red.

B Event that a bus in town in blue.

SR Event that the person sees red.

Given:

$$P(R) = 0.01$$
$$P(B) = 0.99$$
$$P(SR|R) = 0.99$$
$$P(SR|B) = 0.02$$

To find:

$$P(R|SR)$$

Soln :

$$P(R|SR) = \frac{P(SR|R) \cdot P(R)}{P(SR)}$$

Using the law of total probability, we can express $P(SR)$:

$$P(SR) = P(SR|R) \cdot P(R) + P(SR|B) \cdot P(B)$$

Substituting the values:

$$P(SR) = 0.99 \cdot 0.01 + 0.02 \cdot 0.99$$

$$P(SR) = 0.0099 + 0.0198 = 0.0297$$

Now substituting back into the equation for $P(R|SR)$:

$$P(R|SR) = \frac{0.99 \cdot 0.01}{0.0297}$$

$$P(R|SR) = \frac{0.0099}{0.0297}$$

$$P(R|SR) \approx 0.3333$$

Thus, the probability that the bus is red given that the person sees red is approximately 0.3333.

Question 5

Events:

A Event that a voter prefers A.

B Event that a voter prefers B.

Wi Event that exactly i out of 3 people prefer A.

Given:

$$P(A) = 0.95$$

$$P(B) = 0.05$$

Let's say i out of the three people preferred A in the poll and the rest preferred B.

Ways to choose i people from 3:

$$\binom{3}{i} = \frac{3!}{i!(3-i)!} \tag{1}$$

The probability of choosing i people who prefer A and 3-i people who prefer B is given by:

$$P(Wi) = \binom{3}{i} \cdot (0.95)^i \cdot (0.05)^{3-i}$$

Values of i can be 0, 1, 2, or 3. The probability that the exit poll declared majority of A is given by:

$$P(W2) + P(W3)$$

Calculating $P(W2)$:

$$P(W2) = \binom{3}{2} \cdot (0.95)^2 \cdot (0.05)^1$$

$$= 3 \cdot (0.95)^2 \cdot (0.05)$$

$$= 3 \cdot 0.9025 \cdot 0.05$$

$$= 3 \cdot 0.045125$$

$$= 0.135375$$

Calculating $P(W3)$:

$$P(W3) = \binom{3}{3} \cdot (0.95)^3 \cdot (0.05)^0$$

$$= 1 \cdot (0.95)^3 \cdot 1$$

$$= (0.95)^3$$

$$= 0.857375$$

Thus, the total probability that the exit poll declared majority of A is:

$$P(W2) + P(W3) = 0.135375 + 0.857375$$

$$= 0.99275$$

Question 6

Number of people in the town = m
Number of people in the subset = n
Probability that a person prefers A = p

$$q(\mathcal{S}) = \frac{\sum_{i \in I(\mathcal{S})} x_i}{n}$$

where $I(\mathcal{S})$ is a set containing the index (from 1 to m) of each voter in $\mathcal{S}$ and $x_i = 1$ iff i^{th} person prefers A.

Total number of subsets of size n from $m = m^n$

Number of subsets of size n such that i number of people out of them prefer A =

$$m^n \sum_{i=0}^{n} (1-p)^{n-i} (p)^i \binom{n}{i}$$

$q(\mathcal{S})$ for each such subset where i people prefer A is

$$q(\mathcal{S}) = \frac{i}{n}$$

Thus $q(\mathcal{S})$ for all such subset where i people prefer A is

$$m^n \sum_{i=0}^{n} \frac{i}{n} \cdot \binom{n}{i} (1-p)^{n-i} (p)^i$$

(a)

To show:

$$\sum_{\mathcal{S}} \frac{q(\mathcal{S})}{m^n} = p.$$

Soln:
Consider the binomial equation

$$(xp + (1-p))^n = \sum_{i=0}^{n} \binom{n}{i} (1-p)^{n-i} (x)^i (p)^i$$

Differentiating both sides with respect to x:
LHS :

$$\frac{d}{dx}((xp + (1-p))^n) = n(xp + (1-p))^{n-1} p$$

RHS:

$$\frac{d}{dx}(\sum_{i=0}^{n} \binom{n}{i} (1-p)^{n-i} x^i p^i) = \sum_{i=0}^{n} i \cdot \binom{n}{i} (1-p)^{n-i} x^{i-1} p^i$$

Substituting $x = 1$:
LHS:

$$n(p + (1-p))^{n-1} p = n \cdot p$$

RHS:

$$\sum_{i=0}^{n} i \cdot \binom{n}{i} (1-p)^{n-i} p^i$$

Thus, we have:

$$\sum_{i=0}^{n} i \cdot \binom{n}{i} (1-p)^{n-i} p^i = n \cdot p \tag{2}$$

Dividing both sides by n:

$$\sum_{i=0}^{n} \frac{i}{n} \cdot \binom{n}{i} (1-p)^{n-i} p^i = p$$

Now, multiplying both sides by m^n:

$$m^n \sum_{i=0}^{n} \frac{i}{n} \cdot \binom{n}{i} (1-p)^{n-i} p^i = m^n \cdot p$$

Thus, we have:

$$\sum_{\mathcal{S}} \frac{q(\mathcal{S})}{m^n} = p$$

(b)

To show:

$$\sum_{S} \frac{q^2(\mathcal{S})}{m^n} = \frac{p}{n} + \frac{p^2(n-1)}{n}.$$

Soln: $q^2(\mathcal{S})$ for each such subset where i people prefer A is

$$q^2(\mathcal{S}) = (\frac{i}{n})^2$$

Thus $q^2(\mathcal{S})$ for all such subset where i people prefer A is

$$m^n \sum_{i=0}^{n} \frac{i^2}{n^2} \cdot \binom{n}{i} (1-p)^{n-i} (p)^i$$

Using the binomial equation:

$$(xp + (1-p))^n = \sum_{i=0}^{n} \binom{n}{i} (1-p)^{n-i} x^i p^i$$

Differentiating twice with respect to x:
LHS:

$$\frac{d^2}{dx^2}((xp + (1-p))^n) = n(n-1)(xp + (1-p))^{n-2} p^2$$

RHS:

$$\frac{d^2}{dx^2}(\sum_{i=0}^{n} \binom{n}{i} (1-p)^{n-i} x^i p^i) = \sum_{i=0}^{n} i(i-1) \cdot \binom{n}{i} (1-p)^{n-i} x^{i-2} p^i$$

Substituting $x = 1$:
LHS:

$$n(n-1)(p + (1-p))^{n-2} p^2 = n(n-1)p^2$$

RHS:

$$\sum_{i=0}^{n} i(i-1) \cdot \binom{n}{i} (1-p)^{n-i} p^i$$

Thus, we have:

$$\sum_{i=0}^{n} i(i-1) \cdot \binom{n}{i} (1-p)^{n-i} p^i = n(n-1)p^2$$

$$\sum_{i=0}^{n} i^2 \cdot \binom{n}{i} (1-p)^{n-i} p^i - \sum_{i=0}^{n} i \cdot \binom{n}{i} (1-p)^{n-i} p^i = n(n-1)p^2$$

Using 2:

$$\sum_{i=0}^{n} i^2 \cdot \binom{n}{i}(1-p)^{n-i}p^i - np = n(n-1)p^2$$

Dividing both sides by n^2 and nultiplying by m^n:

$$m^n \sum_{i=0}^{n} \frac{i^2}{n^2} \cdot \binom{n}{i}(1-p)^{n-i}p^i - m^n \cdot \frac{p}{n} = m^n \cdot \frac{(n-1)p^2}{n}$$

Thus, we have:

$$\sum_{S} \frac{q^2(\mathcal{S})}{m^n} = \frac{p}{n} + \frac{p^2(n-1)}{n}$$

(c)

To show:

$$\sum_{S} \frac{(q(S)-p)^2}{m^n} = \frac{p(1-p)}{n}.$$

Soln: We can rewrite the left-hand side as:

$$\sum_{S} \frac{q^2(\mathcal{S})}{m^n} + \sum_{S} \frac{p^2}{m^n} - 2p \sum_{S} \frac{q(\mathcal{S})}{m^n}$$

Substituting the value of $\sum_S \frac{q(\mathcal{S})}{m^n} = p$:

$$\sum_{S} \frac{q^2(\mathcal{S})}{m^n} + \sum_{S} \frac{p^2}{m^n} - 2p^2$$

Now substituting the value of $\sum_S \frac{q^2(\mathcal{S})}{m^n} = \frac{p}{n} + \frac{p^2(n-1)}{n}$:

$$\frac{p}{n} + \frac{p^2(n-1)}{n} - 2p^2 + m^n \frac{p^2}{m^n}$$

Hence

$$\sum_{S} \frac{(q(S)-p)^2}{m^n} = \frac{p^2 n + p - p^2}{n} - p^2$$

Thus, we have shown that:

$$\sum_{S} \frac{(q(S)-p)^2}{m^n} = \frac{p(1-p)}{n}$$

(d)

$q_i(\mathcal{S})$: $q(\mathcal{S})$ for i^{th} subset of n elements.

There are m^n subsets of size n from m people.

Let

$$S_k = \{q(\mathcal{S}) : |q(S) - p| > \delta\}$$

Consider σ as the standard deviation of $q(\mathcal{S})$:

$$\sigma = \sqrt{\sum_{\mathcal{S}} \frac{(q_i(\mathcal{S}) - p)^2}{m^n - 1}}$$

We can write

$$\delta = \delta \cdot \frac{\sigma}{\sigma}$$

Let $\frac{\delta}{\sigma} = k$ Clealry $k > 0$. So,

$$S_k = \{q(\mathcal{S}) : |q(S) - p| > k\sigma\}$$

Using Two-sided Chebyshev's inequality:

$$\frac{|S_k|}{m^n} \leq \frac{1}{k^2}$$

$$\implies \frac{|S_k|}{m^n} \leq \frac{\sigma^2}{\delta^2}$$

$$\implies \frac{|S_k|}{m^n} \leq \frac{\left(\sum_{\mathcal{S}} \frac{(q_i(\mathcal{S})-p)^2}{m^n-1}\right)}{\delta^2}$$

Now,

$$\sum_{\mathcal{S}} \frac{(q_i(\mathcal{S}) - p)^2}{m^n - 1} = \sum_{\mathcal{S}} \frac{(q(S) - p)^2}{m^n} \cdot \frac{m^n}{m^n - 1}$$

Using the result from part (c):

$$\sum_{\mathcal{S}} \frac{(q_i(\mathcal{S}) - p)^2}{m^n - 1} = \frac{p(1-p)}{n} \cdot \frac{m^n}{m^n - 1} \leq \frac{p(1-p)}{n}$$

So,

$$\frac{|S_k|}{m^n} \leq \frac{\left(\sum_{\mathcal{S}} \frac{(q_i(\mathcal{S})-p)^2}{m^n-1}\right)}{\delta^2} \leq \frac{p(1-p)}{n} \cdot \frac{1}{\delta^2}$$

This is a very nice application of Chebychev's inequality. Significance of this result is that it gives us a bound on the fraction of subsets whose average preference deviates from the true preference by more than a certain amount, δ. We notice that this proportion is quite small which means we can be confident that most subsets will have an average preference prtty close to the true preference p.