Homework 1

Brian Cervantes Alvarez

May 6, 2024

ST 559 Bayesian Statistics

Question 1

Let,

- E_s as the event a car has the smallest engine.

- E_m as the event a car has the medium-size engine.

- E_l as the event a car has the largest engine.

- F as the event a car fails the emissions test.

We are given,

- $P(E_s) = 0.45, P(E_m) = 0.35, P(E_l) = 0.20$

- $P(F \mid E_s) = 0.10, P(F \mid E_m) = 0.12, P(F \mid E_l) = 0.15$

As a result, we apply the law of total probability:

$$P(F) = P(F \mid E_s)P(E_s) + P(F \mid E_m)P(E_m) + P(F \mid E_l)P(E_l)$$

Plugging in the values:

$$P(F) = (0.10)(0.45) + (0.12)(0.35) + (0.15)(0.20)$$

$$P(F) = 0.045 + 0.042 + 0.03 = 0.117$$

Therefore, the probability that a randomly selected car will fail an emissions test within two years is 11.7%.

Question 2

The Beta distribution for Y is defined by its probability density function:

$$f(y \mid a, b) = \frac{y^{a-1}(1-y)^{b-1}}{B(a,b)}$$

The expected value $E[Y]$ of Y is given by:

$$E[Y] = \int_0^1 y f(y \mid a, b) \, dy$$

Substituting the density function, we have:

$$E[Y] = \int_0^1 y \frac{y^{a-1}(1-y)^{b-1}}{B(a,b)} \, dy$$

This integral simplifies by recognizing it as a property of the Beta function:

$$E[Y] = \frac{a}{a+b}$$

Next, the variance $\text{Var}(Y)$ of Y is given by it's definition,

$$\text{Var}(Y) = E[Y^2] - (E[Y])^2$$

First, compute $E[Y^2]$:

$$E[Y^2] = \int_0^1 y^2 f(y \mid a, b) \, dy = \int_0^1 y^2 \frac{y^{a-1}(1-y)^{b-1}}{B(a,b)} \, dy$$

This becomes:

$$E[Y^2] = \int_0^1 y^{a+1-1}(1-y)^{b-1} \frac{1}{B(a,b)} \, dy$$

Recognizing this as a Beta function integral:

$$E[Y^2] = \frac{B(a+2,b)}{B(a,b)} = \frac{(a+1)a}{(a+b+1)(a+b)}$$

Using $E[Y] = \frac{a}{a+b}$,

$$\text{Var}(Y) = \frac{(a+1)a}{(a+b+1)(a+b)} - \left(\frac{a}{a+b}\right)^2$$

$$\text{Var}(Y) = \frac{ab}{(a+b)^2(a+b+1)}$$

Question 3

Let $N \sim \text{Binomial}(n, p)$ and $X_i \sim \text{Poisson}(\lambda)$ be i.i.d. Define $Y = \sum_{i=1}^{N+1} X_i$. To find $E(Y)$ and $\text{Var}(Y)$, we use the properties of conditional expectation and variance:

$$E(Y) = E(E(Y \mid N))$$

Since $E(X_i) = \lambda$:

$$E(Y \mid N) = (N + 1)\lambda$$

$$E(Y) = E((N + 1)\lambda) = \lambda E(N + 1) = \lambda(n + 1)$$

The variance, using the law of total variance, is:

$$\text{Var}(Y) = E(\text{Var}(Y \mid N)) + \text{Var}(E(Y \mid N))$$

$$\text{Var}(Y \mid N) = (N + 1)\lambda, \text{Var}(E(Y \mid N)) = \lambda^2 \text{Var}(N + 1)$$

$$\text{Var}(Y) = \lambda E(N + 1) + \lambda^2(np + p) = \lambda(n + 1) + \lambda^2(np + p)$$

Question 4

Given that $Y \sim \text{Binomial}(n, \theta)$ and $\theta \sim \text{Beta}(a, b)$, we want to find the posterior distribution of θ given $Y = y$.

We can let our prior be,

$$\theta \sim \text{Beta}(a, b)$$

Given θ, the likelihood of observing $Y = y$ under a binomial distribution is,

$$P(Y = y \mid \theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}$$

Using Bayes' theorem,

$$P(\theta \mid Y = y) \propto P(Y = y \mid \theta) P(\theta)$$

Substituting the expressions:

$$P(\theta \mid Y = y) \propto \theta^y (1 - \theta)^{n-y} \times \theta^{a-1} (1 - \theta)^{b-1}$$

$$P(\theta \mid Y = y) \propto \theta^{y+a-1} (1 - \theta)^{n-y+b-1}$$

Recognizing the form of a Beta distribution,

$$\theta \mid Y = y \sim \text{Beta}(y + a, n - y + b)$$

This tells us that after observing y successes in n trials, the prior belief about θ is updated, increasing our confidence in the true value of θ based on the data observed.

Question 5

We flipped a coin 10 times and observed 2 heads. We will use Beta$(1, 1)$ as our prior, for the probability of observing heads (θ), and find the 95% posterior quantile-based interval for θ to determine if the coin might be biased.

Again, given that our prior is,

$$\theta \sim \text{Beta}(1, 1)$$

We can calculate the likelihood of observing 2 heads out of 10 flips,

$$P(Y = 2 \mid \theta) = \binom{10}{2} \theta^2 (1 - \theta)^8$$

Then, our posterior distribution of θ given $Y = 2$ will be,

$$\theta \mid Y = 2 \sim \text{Beta}(2 + 1, 10 - 2 + 1) = \text{Beta}(3, 9)$$

Lastly, to determine the 95% quantile-based interval for θ, we need to compute the 2.5th and 97.5th percentiles of the $Beta(3, 9)$. I used R to do the calculation.

```r
alpha <- 3
beta <- 9
qbeta(c(0.025, 0.975), alpha, beta)
```

```
[1] 0.06021773 0.51775585
```

Our observations from flipping a coin 10 times show that there's a chance it could be biased towards tails, as the confidence interval is mostly below 0.5. However, because 0.5 is still within the interval, we can't be sure the coin is biased with just these 10 flips. To really figure out if the coin is fair, flipping it more times would be helpful, as this could give us a clearer picture by tightening the interval around the true chance of landing heads.