Homework_1_449

Nicholas Wendt

2025-02-04

Questions: 1.2, 1.3, 1.4, 1.6

Question 1.2

Which scale of measurement is most appropriate for the following variables — nominal or ordinal?

- UK political party preference (Labour, Liberal Democrat, Conservative, other) - Nominal
- Highest educational degree obtained (none, high school, bachelor's, master's, doctorate) - Ordinal
- Patient condition (good, fair, serious, critical) - Ordinal
- Hospital location (London, Boston, Madison, Rochester, Toronto) - Nominal
- Favorite beverage (beer, juice, milk, soft drink, wine, other) - Nominal
- Rating of a movie with 1 to 5 stars, representing (hated it, didn't like it, liked it, really liked it, loved it) - Ordinal

Question 1.3

Each of 100 multiple-choice questions on an exam has four possible answers but one correct response. For each question, a student randomly selects one response as the answer.

Part a Specify the probability distribution of the student's number of correct answers on the exam.

Let X be a random variable that represents the number of questions a student gets correct.

$$X \sim Binomial(n = 100, \pi = 0.25)$$

Part b Based on the mean and standard deviation of that distribution, would it be surprising if the student made at least 50 correct responses? Explain your reasoning.

Yes, it would be surprising if a student gets at least 50 questions correct.

```
prob <- 1 - pbinom(49, size = 100, prob = 0.25)
prob
```

```
## [1] 6.638502e-08
```

Question 1.4

In a particular city, the population proportion π supports an increase in the minimum wage. For a random sample of size 2, let Y = number who support an increase

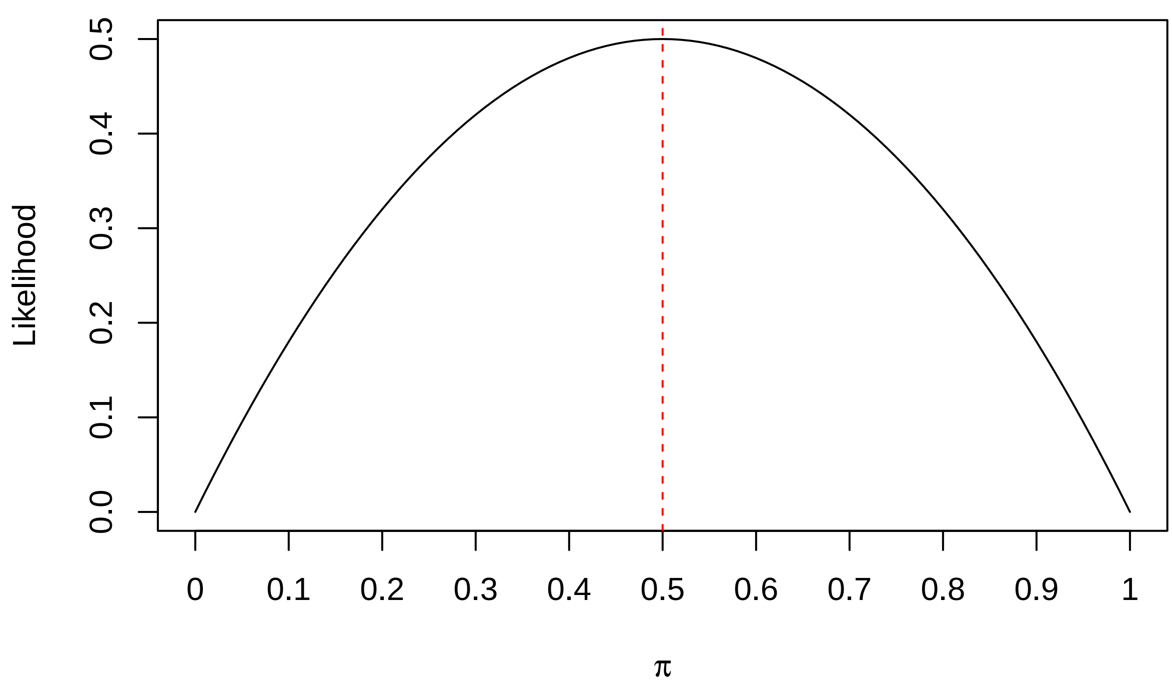
Part a Assuming $\pi = 0.50$, specify the probabilities for the possible values y for Y and find the distribution's mean and standard deviation.

With a sample size of 2, the values y can take on are 0,1,2. The mean of the distribution is $n \cdot \pi = 2 \cdot 0.5 = 1$. The standard deviation of the distribution is $\sqrt{n \cdot \pi(1-\pi)} = \sqrt{2 \cdot 0.5 \cdot 0.5} = \sqrt{5} \approx 0.707$.

Part b Suppose you observe y = 1 and do not know π. Find and sketch the likelihood function. Using the plotted likelihood function, explain why the ML estimate $\hat{\pi} = 0.50$.

$$L(\pi) = \binom{2}{1} \cdot \pi^1 \cdot (1-\pi)^1 = 2\pi(1-\pi) = 2\pi - 2\pi^2$$

```
x = seq(0,1,0.1)
prob = seq(0,1,0.01)
plot(prob, 2*prob - 2*prob^2, type = "l", xaxt = "n", xlab = expression(pi), ylab = paste("Likelihood"))
axis(1, at = x, labels = x)
abline(v = 0.5, col = "red", lty = 2)
```



Since the likelihood function is complicated to derive due to the form of the binomial distribution, we can calculate the derivative of the log-likelihood function.

$$p(y) = \binom{n}{y} \pi^y (1-\pi)^{n-y}$$

$$ln(p(y)) = ln(\binom{n}{y}) + y \cdot ln(\pi) + (n-y) \cdot ln(1-\pi)$$

Now we take the derivative of the function and set it equal to 0 in order to find the maximum

$$\frac{d}{d\pi} ln(p(y)) = 0 + \frac{y}{\hat{\pi}} - \frac{(n-y)}{1-\hat{\pi}} = 0$$

Now solve for $\hat{\pi}$

$$\hat{\pi} = \frac{y}{n}$$

This is a formula used to find the maximum likelihood estimate for π .

In this case, y = 1 and n = 2 making $\hat{\pi} = 0.5$.

The likelihood function looks like a downwards opening parabola or something similar to a normal distribution based on the number of observations and successes. The maximum point of this function at $\frac{y}{n}$.

Question 1.6 Genotypes AA, Aa, and aa occur with probabilities (π_1, π_2, π_3). For n = 3 independent observations, the observed frequencies are (y_1, y_2, y_3).

Part a Explain how you can determine y_3 from knowing y_1 and y_2 . Thus, the multinomial distribution of (y_1, y_2, y_3) is actually two-dimensional.

$y_3 = n - y_2 - y_1$. This is two-dimensional since we only need information of 2 of the 3 y's. y_3 is not independent and is completely determined by y_1 and y_2 .

Part b Show the ten possible observations (y_1, y_2, y_3) with n = 3.

1. $y_1 = 3, y_2 = 0, y_3 = 0$
2. $y_1 = 0, y_2 = 3, y_3 = 0$
3. $y_1 = 0, y_2 = 0, y_3 = 3$
4. $y_1 = 2, y_2 = 1, y_3 = 0$
5. $y_1 = 2, y_2 = 0, y_3 = 1$
6. $y_1 = 1, y_2 = 2, y_3 = 0$
7. $y_1 = 0, y_2 = 2, y_3 = 1$
8. $y_1 = 1, y_2 = 0, y_3 = 2$
9. $y_1 = 0, y_2 = 1, y_3 = 2$
10. $y_1 = 1, y_2 = 1, y_3 = 1$

Part c Suppose (π_1, π_2, π_3) = (0.25, 0.50, 0.25). What probability distribution does y_1 alone have?

$y_1 \sim Binomial(n = n, \pi = 0.25)$. All marginal distributions of a multinomial distribution are Binomial with probability of success $\pi = \pi_i$ and probability of failure $1 - \pi_i = \sum_{\forall j: i \neq j} \pi_j$