Quiz 1 - Statistical Inference

Jean-Luc BELLIER

20 septembre 2016

Question 1

Consider influenza epidemics for two parent heterosexual families. Suppose that the probability is 17% that at least one of the parents has contracted the disease. The probability that the father has contracted influenza is 12% while the probability that both the mother and father have contracted the disease is 6%. What is the probability that the mother has contracted influenza ?

- 17%
- 5%
- 11%
- 6%

Answer : Let call A the probability that the father contracted the disease and B the probability that the mother contracted the disease. We have

$$P(A \cup B) = 0.17$$

$$P(A) = 0.12$$

$$P(A \cap B) = 0.06$$

We also know that

$$P(A \cup B) = 0.17 = P(A) + P(B) - P(A \cap B)$$

So ***P(B) = 11%***.

Question 2

A random variable X is uniform, a box from 0 to 1 of height 1. (So that its density is $f(x) = 1 for 0 \leq x \leq 1$). What is its 75% percentile ? (Hint : look at lecture 2 at 21:30 and chapter 5 problem 5. Also look up the help function for the qunif function in R).

- 0.25
- 0.50
- 0.75
- 0.10

Answer : As the density is constant, ***the point that the area below is 0.75 is also 0.75***.

Question 3

You are playing a game with a friend where you flip a coin and if it comes up heads you give her X dollars and if it comes up tails she gives you Y dollars. The probability that the coin is head is p (some number between 0 and 1). What has to be true between X and Y to make so that both of your total earnings is 0 ? The game would then be "fair".

- $\frac{p}{1-p} = \frac{Y}{X}$
- $X = Y$
- $\frac{p}{1-p} = \frac{X}{Y}$
- $p = \frac{X}{Y}$

Answer : We must have $-p.X + (1-p).Y = 0$. So $\frac{p}{1-p} = \frac{Y}{X}$.

Question 4

A density that looks like a normal density (but may or may not be exactly normal) is exactly symmetric about 0. (Symmetric means if you flip it around 0 it looks the same). What is the median ?

- The median must be 1
- We can't conclude anything about the median
- The median must be 0
- The median must be different from the mean

Answer : The median must be 0 because 50% of the mass is below 0 and 50% above.

Question 5

Consider the following PMF shown below in R :

```r
x <- 1:4
p <- x/sum(x)
temp <- rbind(x,p)
rownames(temp) <- c("X","prob")
temp
```

```
##      [,1] [,2] [,3] [,4]
## X     1.0  2.0  3.0  4.0
## prob  0.1  0.2  0.3  0.4
```

What is the mean ?

- 2
- 3
- 4
- 1

Answer :

```r
sum(p*x)
```

```
## [1] 3
```

So the mean is 3.

Question 6

A web site (http://medicine.ox.ac.uk/bandolier/band64/b64-7.html) for home pregancy tests cites the following :"When the subjects using the test were women who collected and tested their own samples, the overall sensitivity was 75%. Specificity as also low, in the range 52% to 75%." Assume the lower value for the specificity. Suppose a subject has a positive test and that 30% of women taking pregnancy tests are actually pregnant. What number is closest to the probability of pregnancy given a positive test ?

- 10%
- 30%
- 40%
- 20%

Answer : We know that

$$sensitivity = Pr(+|Preg) = 0.75$$

$$specificity = Pr(-|Preg^C) between 0.52 and 0.75$$

$$Pr(Preg) = 0.3$$

Wee want to compute :

$$P(Preg|+) = \frac{P(+|Preg).P(Preg)}{P(+|Preg).P(Preg) + P(+|Preg^C).P(Preg^C)}$$

So

$$P(Preg|+) = \frac{0.75 * 0.3}{0.75 * 0.3 + (1 - 0.52) * 0.7}$$

The probability of pregnancy given a positive test is around 40%.