

# Statistical Inference - Data Science - Quiz 1 - Coursera

## Statistical Inference Quiz 1

This is Quiz 1 from the Statistical Inference course in the Data Science Specialization from John's Hopkins.

### Questions

---

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?

---

- **11%**
- 

```
father <- 12/100
both <- 6/100
either <- 17/100
either - father + both
## [1] 0.11
```

---

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 75th percentile?

---

- **0.75**
- 

```
qunif(p=0.75, min=0, max=1)
## [1] 0.75
```

---

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 heads is p (some number between 0 and 1.) What has to be true about X and Y to make so that both of your expected total earnings is 0. The game would then be called “fair”.

---

- $p/(1-p)=Y/X$
- 

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

---

- **The median must be 0.**
- 

5. Consider the following PMF shown below in 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

- 
- **3**
- 

```
sum(temp[,1]*temp[,2])
## [1] 3
```

---

6. A web site <[www.medicine.ox.ac.uk/bandolier/band64/b64-7.html](http://www.medicine.ox.ac.uk/bandolier/band64/b64-7.html)> for home pregnancy 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 was 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 the positive test?

---

- **Scatterplots with many many points**

---

```
spec <- .52
sens <- .75
prev <- .3
n <- 1000

TP <- (n*prev*sens)
FP <- (n*(1-prev)*(1-spec))
TN <- (n*(1-prev)*spec)
FN <- (n*prev*(1-spec))

TP/(TP+FP)

## [1] 0.4010695
```