

QUIZ 2 ANSWERS

1.  $X$  is known to be normally distributed with mean 0 and unknown variance  $\sigma^2 = \theta$ . From a sample of 5 you get the numbers -10,7,-1,8,12.

The density function for the normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

We are given that  $\mu = 0$  and  $\sigma^2 = \theta$ . So the density function for this problem is

$$f(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/2\theta}$$

(1) (5pts) Find the MLE estimator and estimate for  $\theta$ .

$$\begin{aligned} L(\theta) &= \prod f(X_i) \\ \ln L(\theta) &= \sum \ln f(X_i) = \sum \left[ -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \theta - \frac{X_i^2}{2\theta} \right] \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \theta - \frac{\sum X_i^2}{2\theta} \\ \frac{\partial}{\partial \theta} \ln L(\theta) &= -\frac{n}{2\theta} + \frac{\sum X_i^2}{2\theta^2} \end{aligned}$$

Set this equal to zero and solve for  $\theta$  to get the estimator

$$\hat{\theta}_{MLE} = \frac{1}{n} \sum X_i^2$$

The estimate is 71.6.

(2) (5pts) Is your estimator unbiased? Explain.

Yes, this is the same question that I asked on the practice quiz.

(3) (5pts) Assuming that your estimator is unbiased, find the Cramér-Rao lower bound for its variance.

Now we need the second derivative of the log likelihood:

$$\frac{\partial^2}{\partial \theta^2} \ln L(\theta) = \frac{n}{2\theta^2} - \frac{\sum X_i^2}{\theta^3}$$

The Fisher information is negative the expected value of this quantity:

$$\mathbb{E} \left[ -\frac{\partial^2}{\partial \theta^2} \ln L(\theta) \right] = -\frac{n}{2\theta^2} + \frac{n\theta}{\theta^3} = \frac{n}{2\theta} = nI(\theta)$$

( $I(\theta)$  is what you get when  $n = 1$ ) The Cramér-Rao lower bound for the variance of our estimator is

$$\text{Var}(\hat{\theta}_{MLE}) \geq \frac{1}{nI(\theta)} = \frac{2\theta}{n}$$

2.  $Y$  is taken from a uniform distribution which goes from  $-\theta$  to  $2\theta$ . So its density function is  $f(y) = \frac{1}{3\theta}$  if  $-\theta \leq y \leq 2\theta$  and  $f(y) = 0$  otherwise. Given a sample of 5 you get -5,2,4,-9,17.

(1) (5pts) Find the MOM estimator and estimate for  $\theta$

The expected value of  $X$  is

$$\mathbb{E}(X) = \int_{-\theta}^{2\theta} \frac{x}{3\theta} dx = \frac{x^2}{6\theta} \Big|_{-\theta}^{2\theta} = \frac{\theta}{2}$$

Set this equal to  $\bar{X}$  and solve for  $\theta$  to get the estimator

$$\hat{\theta}_{MOM} = 2\bar{X}$$

The estimate is

$$\hat{\theta}_{MOM} = 2(1.8) = 3.6$$

(2) (5pts) The MLE is the smallest possible value of  $\theta$  given the data. Find  $\hat{\theta}_{MLE}$ .

The smallest possible value of  $\theta$  is  $\hat{\theta}_{MLE} = 9$ . (It is either the largest absolute value of a negative number (9) or half the largest positive number ( $17/2=8.5$ ) whichever is bigger.)

Explain why this is the MLE.

The likelihood of  $\theta$  is

$$L(\theta) = \prod f(X_i) = \left(\frac{1}{3\theta}\right)^n$$

provided that all the numbers  $X_i$  are between  $-\theta$  and  $2\theta$ . Since  $\theta$  is in the denominator, the likelihood is the largest when  $\theta$  is as small as possible.

(3) (5pts) The MLE is sufficient for  $\theta$ . Explain what this means.

It means that, given that  $\hat{\theta} = 9$ , (so that one of the numbers is either -9 or 18), there is no additional information contained in the other 4 numbers. Mathematically speaking, the other four numbers are uniformly distributed between  $-9$  and  $18$ .

3. (extra credit)

(5pts) In the previous problem, the MOM is unbiased and the MLE is sufficient. Can you combine them using the Rao-Blackwell formula?

The Rao-Blackwell formula gives a new estimator  $\hat{\theta}^*$  which is both unbiased and sufficient:

$$\hat{\theta}^* = \mathbb{E}(\hat{\theta}_{MOM} | \hat{\theta}_{MLE})$$

Let's first use the numbers. Given that the MLE estimate is  $\hat{\theta}_{MLE} = 9$ , the other four numbers are somewhere between  $-9$  and  $18$  with uniform density. So, their expected value is  $9/2 = 4.5$ . Each number is twice as likely to be positive as negative and this also applies to the extreme number: it is twice as likely to be  $18$  as it is to be  $-9$ . So its expected value is

$$\frac{1}{3}(-9) + \frac{2}{3}(18) = 9$$

So

$$\mathbb{E}(2\bar{X}|\hat{\theta}_{MLE}) = \frac{2}{n} [(n-1)4.5 + 9] = 10.8$$

This is the estimate. The estimator is

$$\mathbb{E}(2\bar{X}|\hat{\theta}_{MLE}) = \frac{2}{n} \left[ (n-1) \frac{\hat{\theta}_{MLE}}{2} + \hat{\theta}_{MLE} \right] = \frac{n+1}{n} \hat{\theta}_{MLE}$$

(5pts) Find the 95% confidence interval for  $\sigma^2$  in the first problem.

The expected value of  $X^2$  is  $\sigma^2 + \mu^2 = \sigma^2 = \theta$ . Its variance is

$$\text{Var}(X^2) = \mathbb{E}(X^4) - \mathbb{E}(X^2)^2 = 3\sigma^4 - \sigma^4 = 2\sigma^4 = 2\theta^2$$

because:

$$\begin{aligned} \mathbb{E}(X^4) &= \int_{-\infty}^{\infty} x^4 f(x) dx = \int_{-\infty}^{\infty} (-x^3 \theta) df(x) \\ &= (-x^3 \theta) f(x) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (-3x^2 \theta) f(x) dx = 3\theta \mathbb{E}(X^2) = 3\theta \sigma^2 = 3\sigma^4 = 3\theta^2 \end{aligned}$$

So, if we use the normal approximation (the true distribution is  $\chi^2$ ), we get that the 95% probability statement

$$\begin{aligned} \mathbb{P} \left( \sigma^2 - 1.96 \frac{1}{\sqrt{n}} \sqrt{2}\sigma^2 \leq \frac{1}{n} \sum X_i^2 \leq \sigma^2 + 1.96 \frac{1}{\sqrt{n}} \sqrt{2}\sigma^2 \right) &= .95 \\ \mathbb{P}(-0.24\sigma^2 \leq \frac{1}{5} \sum X_i^2 \leq 2.23\sigma^2) &= .95 \end{aligned}$$

Since a sum of squares cannot be negative, we should ignore the inequality on the left. This gives the 95% confidence statement

$$\frac{71.6}{2.23} \leq \sigma^2$$

The 95% confidence interval for  $\sigma^2$  is

$$[31.97, \infty)$$

When we learn about  $\chi^2$  we can come back to this problem to see how accurate this answer is.