## **Multiple Linear Regression**

Inference for the MLE

Math 392

#### Running example

Let

$$X \sim \operatorname{Pois}(\theta)$$

$$f(x| heta)=rac{1}{x!} heta^x e^{- heta},\quad x\geq 0,\quad heta>0,\quad \hat{ heta}^{MLE}=ar{x}$$

The likelihood function *L*:

$$L(\theta) = f(x_1|\theta)f(x_2|\theta)\dots f(x_n|\theta)$$

The log-likelihood function *l*:

$$l( heta) = \log(f(x_1| heta)) + \log(f(x_1| heta)) + \ldots + \log(f(x_1| heta))$$

The  $\hat{\theta}^{MLE}$  is the solution to setting  $\frac{\partial}{\partial \theta} \log(f(\mathbf{x}|\theta))$  to 0.

#### **Score function**

Define

$$U = rac{\partial}{\partial heta} \mathrm{log}(f(X| heta))$$

as the score function or score statistic.

#### Note:

- $\theta$  is fixed
- *X* is a single RV
- therefore U is a RV

## Understanding $oldsymbol{U}$

What is its *expected value*?

$$E(U) = E(\frac{\partial}{\partial \theta} \log(f(x|\theta)))$$

$$= \int \frac{\partial}{\partial \theta} \log(f(x|\theta)) f(x|\theta) dx$$

$$= \int \frac{\partial}{\partial \theta} f(x|\theta) \frac{1}{f(x|\theta)} f(x|\theta) dx$$

$$= \int \frac{\partial}{\partial \theta} f(x|\theta) dx$$

$$= \frac{\partial}{\partial \theta} \int f(x|\theta) dx$$

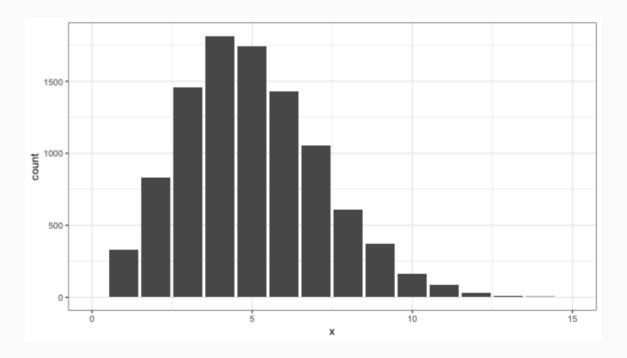
$$= \frac{\partial}{\partial \theta} \int 1$$

$$= 0$$

#### **Poisson RV**

Let  $X \sim \text{Poi}(\theta)$ .

```
n <- 10000
theta <- 5
x <- rpois(n, theta)
```



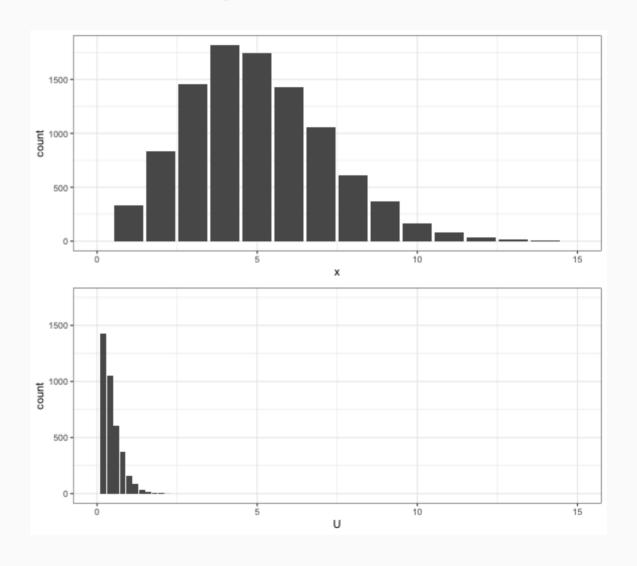
#### Score RV

Define 
$$U = \frac{\partial}{\partial \theta} \log f(X|\theta)$$
.

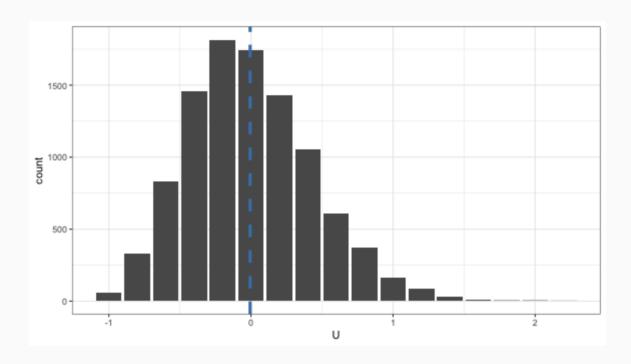
$$egin{aligned} U &= rac{\partial}{\partial heta} \log rac{1}{x!} heta^x e^{- heta} \ &= rac{\partial}{\partial heta} (-\log(x!) + x \log( heta) - heta) \ &= 0 + x rac{1}{ heta} - 1 \end{aligned}$$

U <- x / theta - 1

## Distribution of X vs U



## Distribution of $oldsymbol{U}$



### Finding the variance of $oldsymbol{U}$

Recall  $Var(U) = E(U^2) - E(U)^2$ , so we seek  $E(U^2)$ . Begin by writing down our previous result, that the expected value is zero, and take derivatives of both sides.

$$\begin{split} \frac{\partial}{\partial \theta} 0 &= \frac{\partial}{\partial \theta} \int \frac{\partial}{\partial \theta} \log(f(x|\theta)) f(x|\theta) \mathrm{d}x \\ 0 &= \int \frac{\partial^2}{\partial \theta^2} \log(f(x|\theta)) f(x|\theta) \mathrm{d}x + \int \frac{\partial}{\partial \theta} \log(f(x|\theta)) \frac{\partial}{\partial \theta} f(x|\theta) \mathrm{d}x \\ 0 &= E\left(\frac{\partial^2}{\partial \theta^2} \log(f(x|\theta))\right) + \int \frac{\partial}{\partial \theta} \log(f(x|\theta)) \frac{\frac{\partial}{\partial \theta} f(x|\theta)}{f(x|\theta)} f(x|\theta) \mathrm{d}x \\ 0 &= E\left(\frac{\partial^2}{\partial \theta^2} \log(f(x|\theta))\right) + \int \frac{\partial}{\partial \theta} \log(f(x|\theta)) \frac{\partial}{\partial \theta} \log(f(x|\theta)) f(x|\theta) \mathrm{d}x \end{split}$$

$$A o Var(U) = E(U^2) = -E\left(rac{\partial^2}{\partial heta^2}{
m log}(f(x| heta))
ight) = I( heta)$$

#### **Score variance**

For  $X \sim \text{Pois}(\theta)$ ,

$$egin{align} Var(U) &= -E\left(rac{\partial^2}{\partial heta^2}\mathrm{log}(f(X| heta))
ight) \ &= -E\left(rac{\partial}{\partial heta}(xrac{1}{ heta}-1)
ight) \ &= -E\left(-xrac{1}{ heta^2}
ight) = rac{1}{ heta}. \end{split}$$

```
1/theta
```

**##** [1] 0.2

var(U)

**##** [1] 0.1961535

### A CLT for $ar{U}$

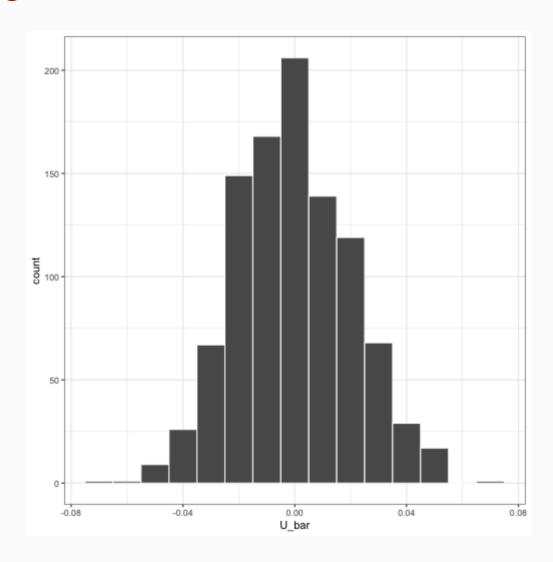
U(X) is a function of a single RV. If we have an iid sample of size n -  $X_1, X_2, \ldots, X_n$  - we have a corresponding iid sample  $U(X_1), U(X_2), \ldots, U(X_n)$ , each with mean 0 and variance  $I(\theta)$ , by the CLT,

$$rac{rac{1}{n}\sum_{i=1}^n U_i - 0}{\sqrt{I( heta)/n}} \quad \stackrel{D}{
ightarrow} \quad N(0,1)$$

## CLT for $ar{U}$

```
n <- 500
theta <- 5
it <- 1000
U_bar <- rep(NA, it)
for (i in 1:it) {
    x <- rpois(n, theta)
    U <- x / theta - 1
    U_bar[i] <- mean(U)
}</pre>
```

## CLT for $ar{U}$



## CLT for $ar{U}$

```
var(U_bar)

## [1] 0.0004181934

(1 / theta) / n

## [1] 4e-04
```

#### The Asymptotic Normality of the MLE

**Theorem**: Let  $X_1, X_2, ..., X_n$  be an iid sample from a regular family with parameter  $\theta$ . Let  $\hat{\theta}^{MLE}$  be the solution to the equation

$$rac{\partial}{\partial heta} ext{log}(f(x1,x2,\ldots,x_n| heta)) = 0$$

then

$$\sqrt{nI( heta)}(\hat{ heta}^{MLE}- heta) \quad \stackrel{D}{
ightarrow} \quad N(0,1)$$

**Proof**: Consider  $U_i = \frac{\partial}{\partial \theta} \log(f(X_i|\theta))$  as a function of both  $X_i$  and  $\theta$ . If we sum the  $U_i$  and expand around the true value:

$$\sum_{i=1}^n rac{\partial}{\partial heta} \mathrm{log}(f(x_i|\hat{ heta})) pprox \sum_{i=1}^n rac{\partial}{\partial heta} \mathrm{log}(f(x_i| heta)) + \left[\sum_{i=1}^n rac{\partial^2}{\partial heta^2} \mathrm{log}(f(x_i| heta))
ight] (\hat{ heta} - heta)$$

Since  $\hat{\theta}$  is the MLE, the term on the left is 0. Now we can write this as a function of  $U_i$  and rearrange:

$$egin{aligned} \sum_{i=1}^n U_i &= -\sum_{i=1}^n rac{\partial^2}{\partial heta^2} \mathrm{log}(f(x_i| heta))(\hat{ heta} - heta) \ &rac{rac{1}{n} \sum_{i=1}^n U_i}{\sqrt{I( heta)/n}} &= rac{-rac{1}{n} \sum_{i=1}^n rac{\partial^2}{\partial heta^2} \mathrm{log}(f(x_i| heta))}{\sqrt{I( heta)/n}} (\hat{ heta} - heta) \end{aligned}$$

We recognize the LHS as an RV that converges to the standard normal, so what is the RHS?

By the LLN,

$$-rac{1}{n}\sum_{i=1}^{n}rac{\partial^{2}}{\partial heta^{2}}\mathrm{log}(f(x_{i}| heta))
ightarrow I( heta),$$

so we can rewrite the RHS as

$$rac{I( heta)}{(I( heta)/n)^{1/2}}(\hat{ heta}- heta) = \sqrt{nI( heta)}(\hat{ heta}- heta) \quad \stackrel{D}{
ightarrow} \quad N(0,1)$$

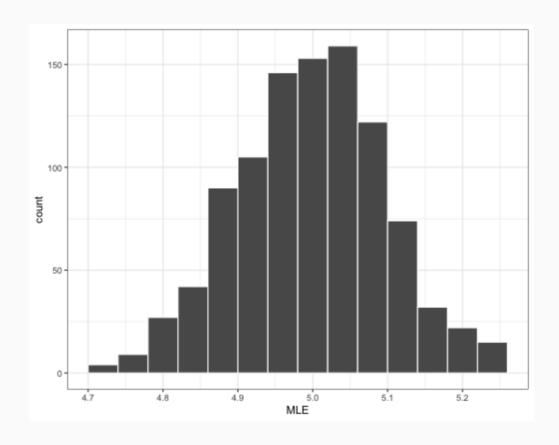
or

$$\hat{ heta} \quad \stackrel{D}{ o} \quad N\left( heta, rac{1}{\sqrt{nI( heta)}}
ight)$$

## CLT for $\hat{ heta}^{MLE}$

```
it <- 1000
MLE <- rep(NA, it)
for (i in 1:it) {
   n <- 500
   theta <- 5
   x <- rpois(n, theta)
   MLE[i] <- mean(x)
}</pre>
```

# CLT for $\hat{ heta}^{MLE}$



## CLT for $\hat{ heta}^{MLE}$

 $z \leftarrow sqrt(n * (1 / theta)) * (MLE - theta)$ 

