Maximum Likelihood

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} F_\theta,\ \theta \in \Theta$$

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

$$\ell(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i; \theta)$$

Close your eyes and differentiate?

Let $X_1, \ldots, X_n$ be a random sample from a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} e^{-x/\beta} x^{\alpha-1}$$

$$\Theta = \{(\alpha, \beta) : \alpha > 0, \beta > 0\}$$

Simulate Some Data: True α=2, β=3

```
> set.seed(3201); alpha=2; beta=3
> D <- round(rgamma(50,shape=alpha, scale=beta),2); D
 [1] 20.87 13.74  5.13  2.76  4.73  2.66 11.74  0.75 22.07 10.49  7.26  5.82 13.08
[14]  1.79  4.57  1.40  1.13  6.84  3.21  0.38 11.24  1.72  4.69  1.96  7.87  8.49
[27]  5.31  3.40  5.24  1.64  7.17  9.60  6.97 10.87  5.23  5.53 15.80  6.40 11.25
[40]  4.91 12.05  5.44 12.62  1.81  2.70  3.03  4.09 12.29  3.23 10.94
> mean(D); alpha*beta
[1] 6.8782
[1] 6
> var(D); alpha*beta^2
[1] 24.90303
[1] 18
```

Alternatives for getting the data into D might be

D = scan(“Gamma.data”)

D = c(20.87, 13.74, …, 10.94)

Log Likelihood

$$\begin{aligned}
\ell(\alpha, \beta) &= \ln \prod_{i=1}^{n} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} e^{-x_i/\beta} x_i^{\alpha-1} \\
&= \ln \left[\beta^{-n\alpha}\, \Gamma(\alpha)^{-n} \exp(-\frac{1}{\beta} \sum_{i=1}^{n} x_i) \left(\prod_{i=1}^{n} x_i \right)^{\alpha-1} \right] \\
&= -n\alpha \ln \beta - \ln \Gamma(\alpha) - \frac{1}{\beta} \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{n} \ln x_i
\end{aligned}$$

R function for the minus log likelihood

$$\ell(\alpha, \beta) = -n\alpha \ln \beta - \ln \Gamma(\alpha) - \frac{1}{\beta} \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{n} \ln x_i$$

```
> # Gamma minus log likelihood: alpha=a, beta=b
> gmll <- function(theta,datta)
+      {
+       a <- theta[1]; b <- theta[2]
+       n <- length(datta); sumd <- sum(datta)
+       sumlogd <- sum(log(datta))
+       gmll <- n*a*log(b) + n*lgamma(a) + sumd/b - (a-1)*sumlogd
+      gmll
+      } # End function gmll
```

Where should the numerical search start?

- How about Method of Moments estimates?
- $E(X) = \alpha\beta$, $Var(X) = \alpha\beta^2$
- Replace population moments by sample moments and put a ~ above the parameters.

$$\tilde{\alpha} = \frac{\overline{X}^2}{S^2} \quad \text{and} \quad \tilde{\beta} = \frac{S^2}{\overline{X}}$$

$$\tilde{\alpha} = \frac{\overline{X}^2}{S^2} \quad \text{and} \quad \tilde{\beta} = \frac{S^2}{\overline{X}}$$

```
> momalpha <- mean(D)^2/var(D); momalpha
[1] 1.899754
> mombeta <- var(D)/mean(D); mombeta
[1] 3.620574
```

```
> gammasearch = nlm(gmll,c(momalpha,mombeta),hessian=T,datta=D); gammasearch
$minimum
[1] 142.0316
```

$$-\ell(\widehat{\alpha}, \widehat{\beta})$$

```
$estimate
[1] 1.805930 3.808674
```

$$\widehat{\alpha} = 1.805930 \qquad \widehat{\beta} = 3.808674$$

```
$gradient
[1] 2.847002e-05 9.133932e-06
```

$$\left(-\frac{\partial \ell}{\partial \alpha}, -\frac{\partial \ell}{\partial \beta}\right)'$$

```
$hessian
         [,1]      [,2]
[1,] 36.68932 13.127271
[2,] 13.12727  6.222282
```

$$\mathbf{H} = \left[\frac{\partial^2(-\ell)}{\partial \theta_i \partial \theta_j}\right]$$

```
$code
[1] 1

$iterations
[1] 6

> det(gammasearch$hessian)
[1] 55.96605
```

Likelihood Ratio Tests

$$X_1, \ldots, X_N \overset{i.i.d.}{\sim} F_\theta,\ \theta \in \Theta,$$
$$H_0 : \theta \in \Theta_0 \text{ v.s. } H_A : \theta \in \Theta \cap \Theta_0^c,$$

$$G^2 = -2 \ln \left(\frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} \right)$$

Under H_0, G^2 has an approximate chi-square distribution for large N. Degrees of freedom = number of (non-redundant, linear) equalities specified by H_0. Reject when G^2 is large.

Degrees of Freedom

Express H_0 as a set of linear combinations of the parameters, set equal to constants (usually zeros).

Degrees of freedom = number of *non-redundant* linear combinations (meaning linearly independent).

Suppose $\boldsymbol{\theta} = (\theta_1, \ldots \theta_7)$, with
$H_0 : \theta_1 = \theta_2,\, \theta_6 = \theta_7,\, \frac{1}{3}\left(\theta_1 + \theta_2 + \theta_3\right) = \frac{1}{3}\left(\theta_4 + \theta_5 + \theta_6\right)$

df=3