EMSE 6035: Marketing Analytics for Design Decisions

Uncertainty

John Paul Helveston, Ph.D.
Assistant Professor
Engineering Management & Systems Engineering
The George Washington University

Background: Estimating Utility Model Coefficients Using Maximum Likelihood Estimation

$$\tilde{u}_j = \boldsymbol{\beta}'\mathbf{x}_j + \tilde{\varepsilon}_j$$

$$= \boxed{\beta_1} x_{j1} + \boxed{\beta_2} x_{j2} + \ldots + \tilde{\varepsilon}_j$$

Weights that denote the *relative* value of attributes $x_{j1}, x_{j2}, \ldots$

Estimate β_1 , β_2 , … , by minimizing the negative log-likelihood function:

$$\text{minimize} - \ln(\mathcal{L}) = -\sum_{j=1}^{J} y_j \ln[P_j(\boldsymbol{\beta}|\mathbf{x})]$$

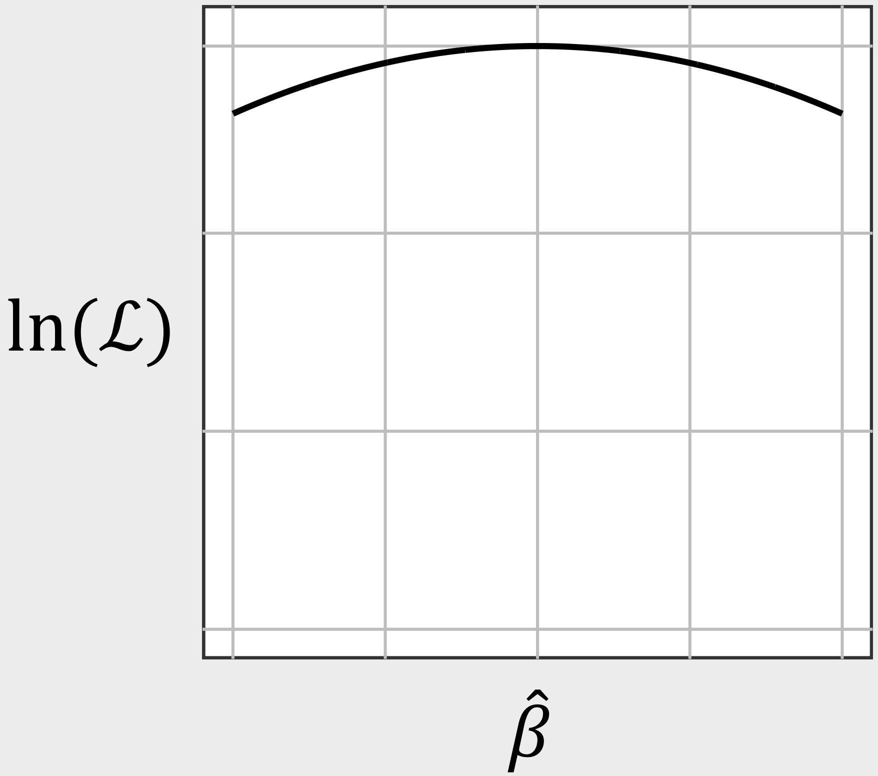
with respect to $\boldsymbol{\beta}$

$y_j = 1$ if alternative j was chosen
$y_j = 0$ if alternative j was not chosen

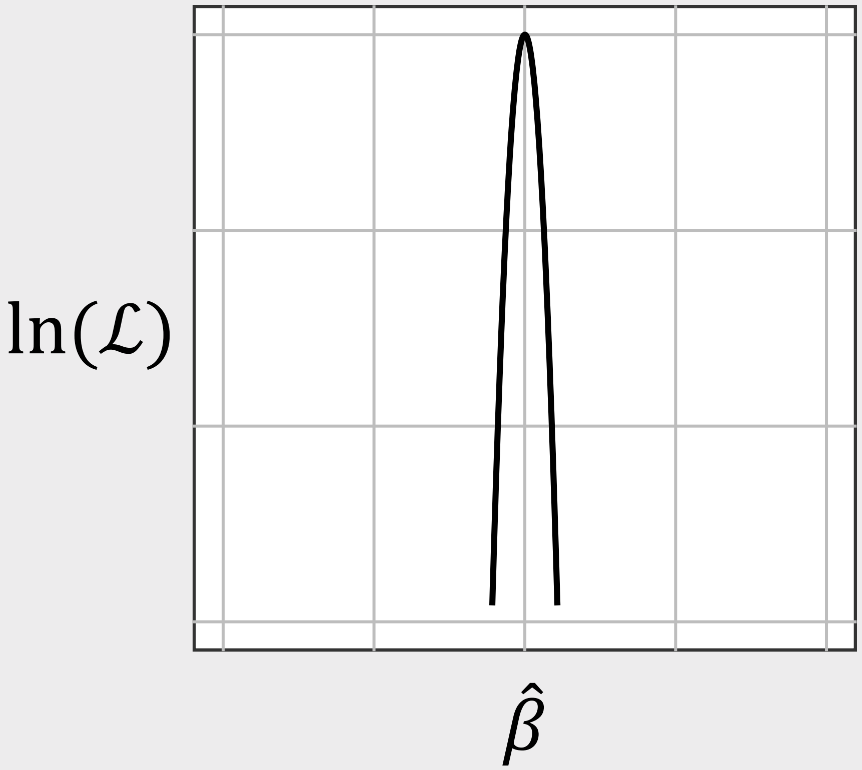
→ Produces point estimates: $\hat{\boldsymbol{\beta}}$
…but these estimates are not precisely known

The certainty of $\widehat{\boldsymbol{\beta}}$ is inversely related to the curvature of the log-likelihood function

Greater variance in $\ln(\mathcal{L})$,
Less certainty in $\hat{\beta}$



Less variance in $\ln(\mathcal{L})$,
Greater certainty in $\hat{\beta}$



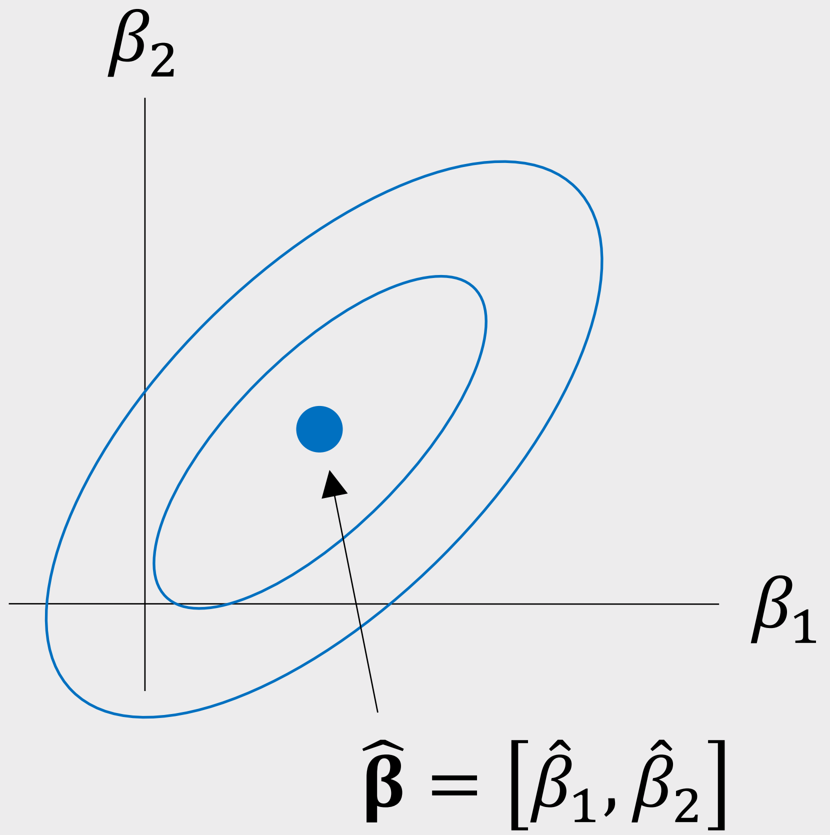
The certainty of $\hat{\boldsymbol{\beta}}$ is inversely related to the curvature of the log-likelihood function

$$\Sigma_{\boldsymbol{\beta}} = -\overbrace{\left[\nabla^2_{\boldsymbol{\beta}} \ln(\mathcal{L})\right]}^{\text{Hessian}}{}^{-1} = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{m1}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{1n}^2 & \cdots & \sigma_{mn}^2 \end{bmatrix}$$

Covariance of $\hat{\boldsymbol{\beta}}$

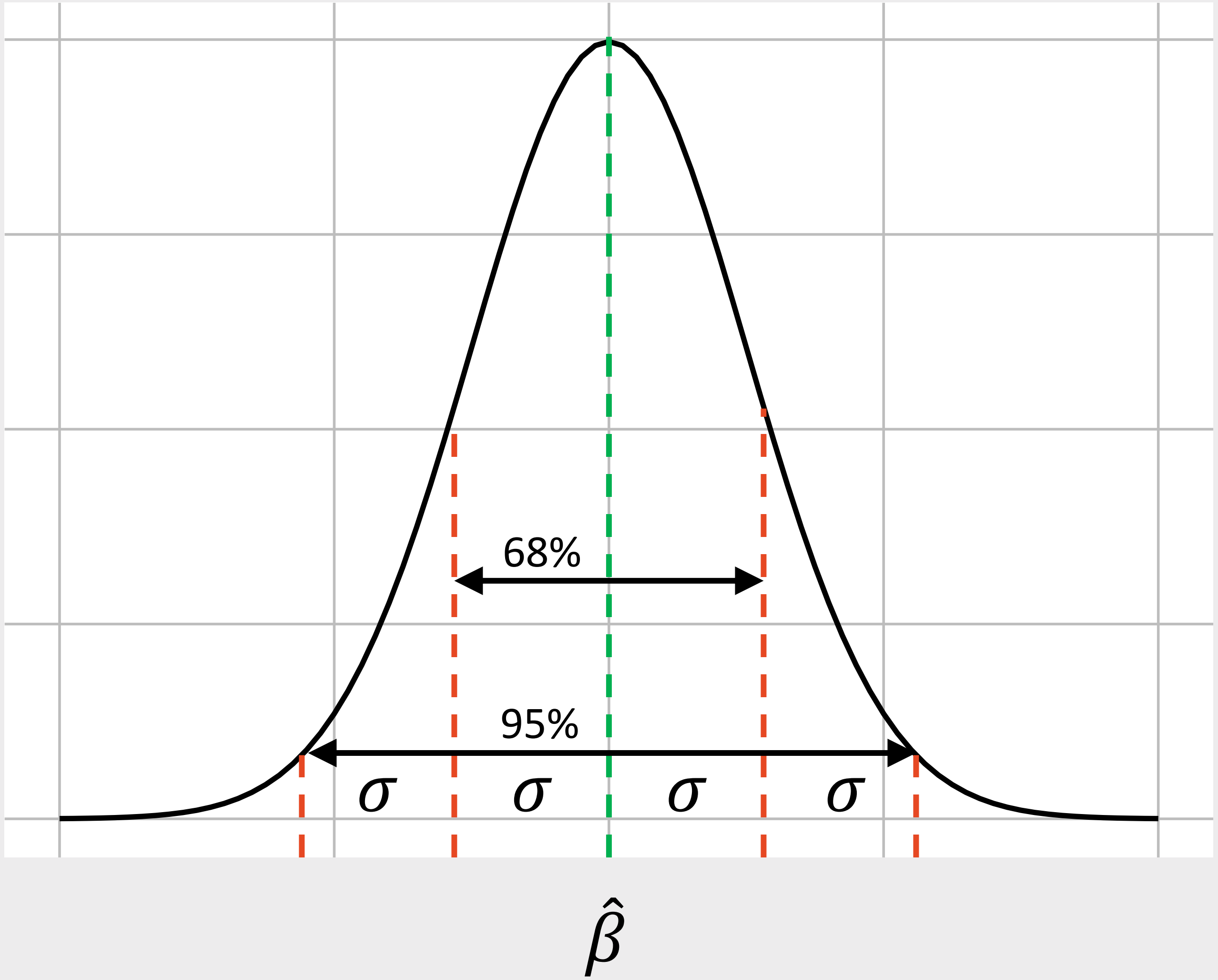
It is common to report $\hat{\boldsymbol{\beta}}$ with its standard errors:

Est.	Std. Err.
$\hat{\beta}_1$	σ_1
$\hat{\beta}_2$	σ_2
$\vdots$	$\vdots$
$\hat{\beta}_m$	σ_m



We use standard errors to report uncertainty about $\hat{\boldsymbol{\beta}}$

Est.	Std. Err.
$\hat{\beta}_1$	σ_1
$\hat{\beta}_2$	σ_2
$\vdots$	$\vdots$
$\hat{\beta}_m$	σ_m



A 95% confidence interval is approximately $[\hat{\beta} - 2\sigma, \hat{\beta} + 2\sigma]$

Practice Question 1

Suppose we estimate a model and get the following results:

$$\widehat{\boldsymbol{\beta}} = [-0.4, 0.5\,] \qquad \nabla^2_{\boldsymbol{\beta}} \ln(\mathcal{L}) = \begin{bmatrix} -6000 & 60 \\ 60 & -700 \end{bmatrix}$$

a) Use the hessian to compute the standard errors for $\widehat{\boldsymbol{\beta}}$.

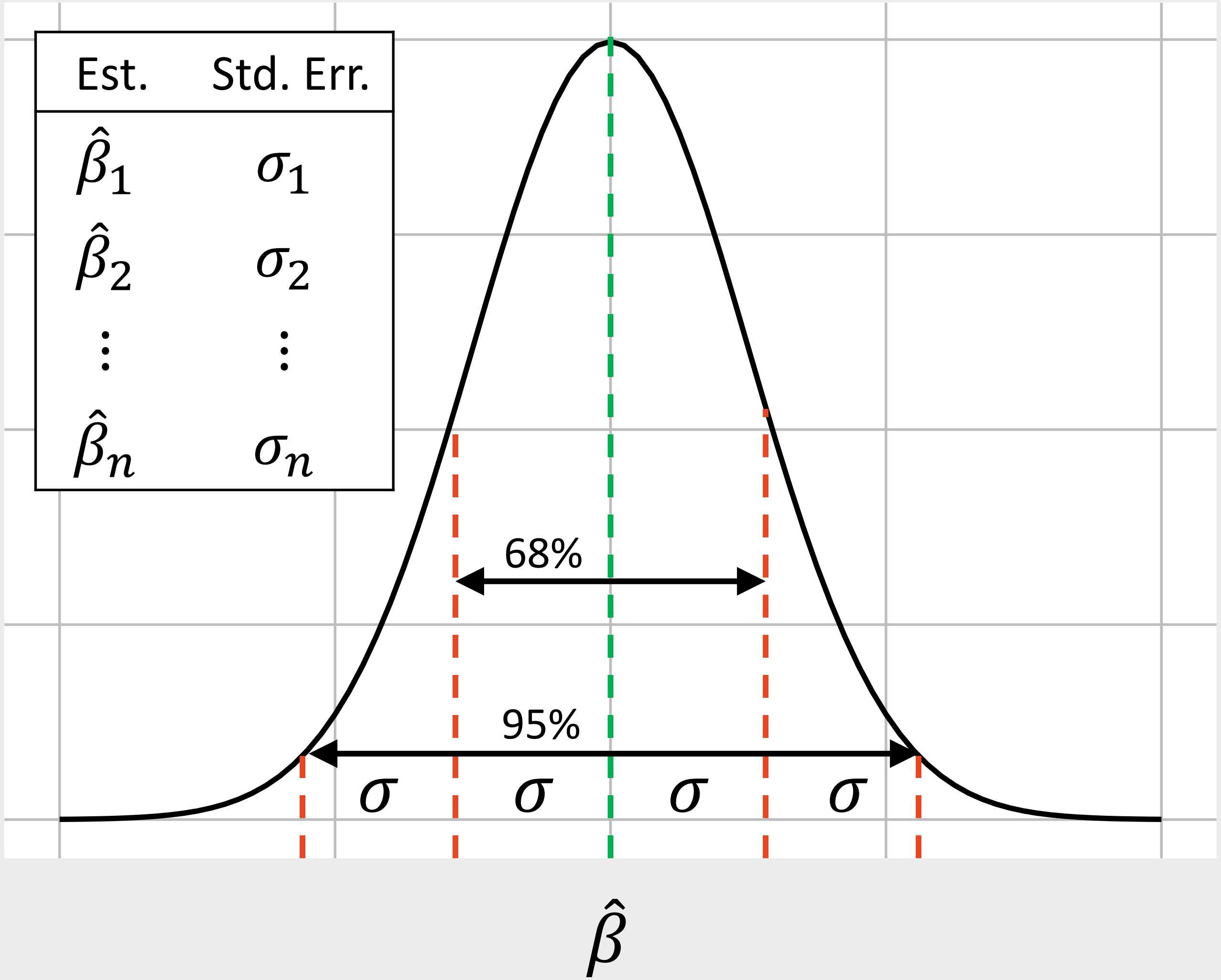
b) Use the standard errors to compute a 95% confidence interval around $\widehat{\boldsymbol{\beta}}$.

Hints:

1. The covariance matrix is computed as $-\left[\nabla^2_{\boldsymbol{\beta}} \ln(\mathcal{L})\right]^{-1}$
2. Use the `matrix()` function to construct a matrix in *R*.
3. Use the `solve()` function to compute the inverse of a matrix in *R*.
4. Use the `diag()` function to get the numbers along the diagonal of a matrix in *R*.

Computing uncertainty via simulation

Use $\hat{\beta}$ and σ to generate samples of $N(\hat{\beta}, \sigma)$



Take sample draws of $\hat{\beta}$ to simulate uncertainty

Example in *R*:

```
> beta = 0.5
> sigma = 0.1
> draws = rnorm(10^5, beta, sigma)
>
> mean(draws)
[1] 0.4996797
> sd(draws)
[1] 0.1001574
>
> c(beta - 2*sigma, beta + 2*sigma)
[1] 0.3 0.7
> quantile(draws, c(0.025, 0.975))
     2.5%     97.5%
0.3044208 0.6964306
```

Sampling $\widehat{\boldsymbol{\beta}}$

$$\boldsymbol{\beta} \sim N\left(\widehat{\boldsymbol{\beta}}, \Sigma\right)$$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \end{bmatrix} \qquad \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1n}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{1n}^2 & \cdots & \sigma_{jn}^2 \end{bmatrix}$$

$$-\Big[\underbrace{\nabla_{\boldsymbol{\beta}}^2 \ln(\mathcal{L})}_{\text{Hessian}}\Big]^{-1}$$

Example in *R*:

```
> library(MASS)
>
> beta = c(price = -0.7, mpg = 0.1, elec=-4.0)
> hessian = matrix(c(
+     -6000,   50,   60,
+        50, -700,   50,
+        60,   50, -300),
+     ncol=3, byrow=T)
> covariance = -1*(solve(hessian))
> draws  = mvrnorm(10^5, beta, covariance)
> head(draws)
          price        mpg      elec
[1,] -0.7184210 0.18428285 -3.951629
[2,] -0.6999711 0.16873388 -3.918036
[3,] -0.7192076 0.11657494 -3.971442
[4,] -0.6851790 0.10707172 -4.039762
[5,] -0.7048889 0.14175661 -4.050028
[6,] -0.6917784 0.09615243 -4.083626
```

Sampling $\widehat{\boldsymbol{\beta}}$

$$\boldsymbol{\beta} \sim N\left(\widehat{\boldsymbol{\beta}}, \Sigma\right)$$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \end{bmatrix} \qquad \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1n}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{1n}^2 & \cdots & \sigma_{jn}^2 \end{bmatrix}$$

$$-\left[\nabla_{\boldsymbol{\beta}}^2 \ln(\mathcal{L})\right]^{-1}$$

Hessian

Example in *R*:

```
> draws <- as.data.frame(draws)
>
> mean(draws$price)
[1] -0.6999623
>
> quantile(draws$price, c(0.025, 0.975))
 2.5%       97.5%
-0.7252586 -0.6747513
```

Practice Question 2

Suppose we estimate the following utility model describing preferences for cars:

$$\tilde{u}_j = \alpha p_j + \beta_1 x_j^{\text{mpg}} + \beta_2 x_j^{\text{elec}} + \varepsilon_j$$

The estimated model produces the following coefficients:

Parameter	Coef.
α	-0.7
β_1	0.1
β_2	-4.0

Hessian		
-6000	50	60
50	-700	50
60	50	-300

a) Generate 10,000 draws of the model coefficients using the estimated coefficients and hessian. Use the `mvrnorm()` function from the `MASS` library.

b) Use the draws to compute the mean and 95% confidence intervals of each parameter estimate.