Standard Errors

Linear and Generalized Linear Models

Learning Objectives

- Standard Errors
  - Linear Regression
  - GLM
- Sampling Distributions

Linear Regression

Standard Errors

- Find the variance of the estimate
- Find the information matrix
- Use for Inference

Finding the Variance

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$Y_i \overset{iid}{\sim} N(\beta_0 + \beta_1 x, \sigma^2)$$

$$\bar{Y} = \frac{1}{n} \sum Y_i$$

$$\bar{Y} \sim N\left(\beta_0 + \beta_1 x, \frac{\sigma^2}{n}\right)$$

$$Var(\hat{\beta}_0) = Var(\bar{Y} - \hat{\beta}_1 \bar{X})$$

$$= Var(\bar{Y}) + \bar{X}^2 \, Var(\hat{\beta}_1)$$

$$\frac{\sigma^2}{n} + \bar{X}^2 \, Var(\hat{\beta}_1)$$

$$Var(\hat{\beta}_1)$$

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\frac{1}{\left(\sum (x_i - \bar{x})^2\right)^2} \sum (x_i - \bar{x})^2 \, Var(y_i - \bar{y})$$

$$Var(\hat{\beta})$$

$$SE(\hat{\beta}) = \sqrt{Var(\hat{\beta})}$$

Estimate for σ^2

$Y = \beta_0 + \beta_1 X + \varepsilon$

$\varepsilon \sim N(0, \sigma^2)$

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} (Y_i - \boldsymbol{X}_i^{\mathrm{T}} \hat{\boldsymbol{\beta}})^2$$

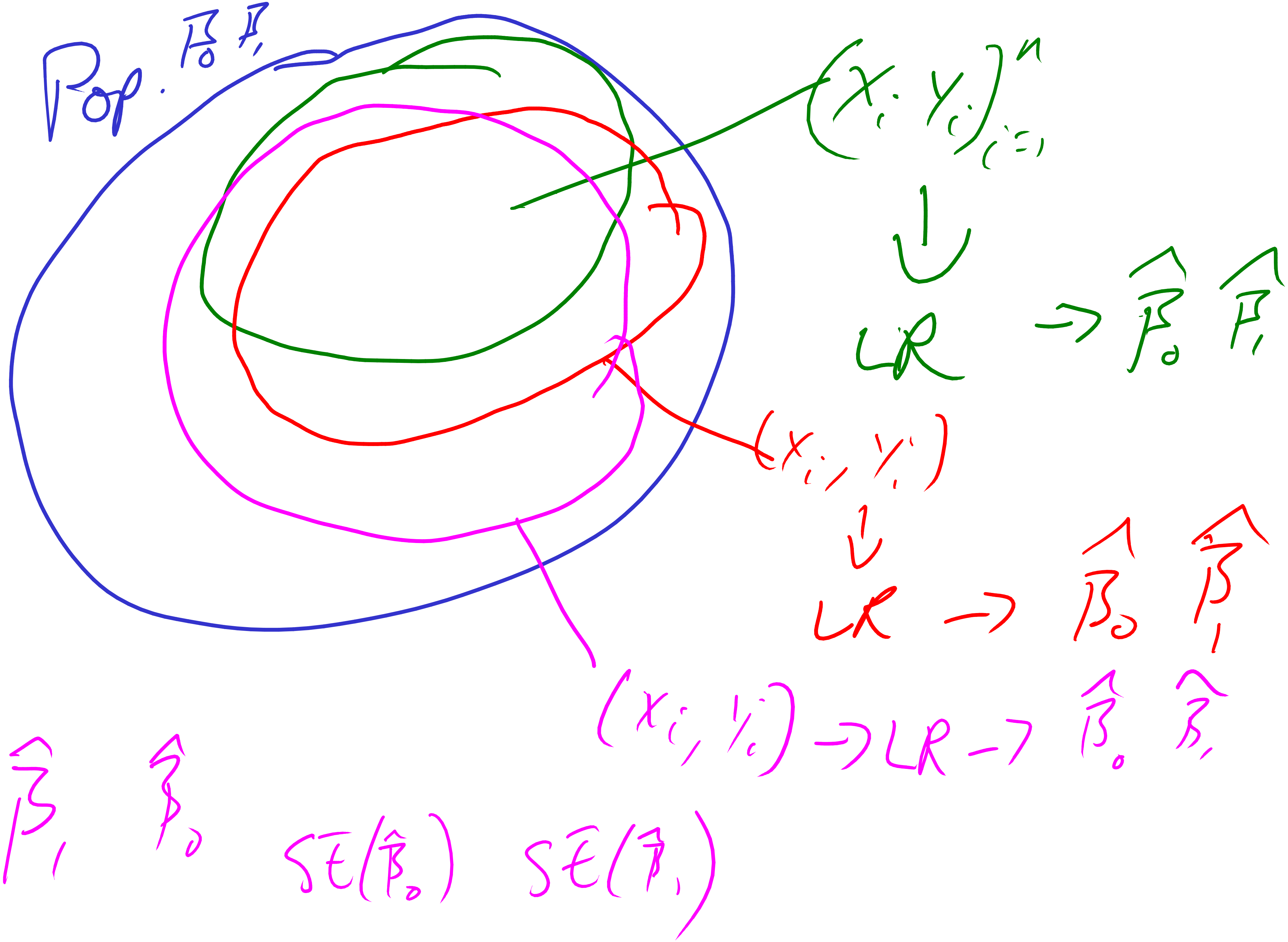
$\hat{\sigma}^2$: Model error

Standard Errors of β's

$$SE(\hat{\beta}_0) = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 \hat{\sigma}^2}{n \sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

$$SE(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

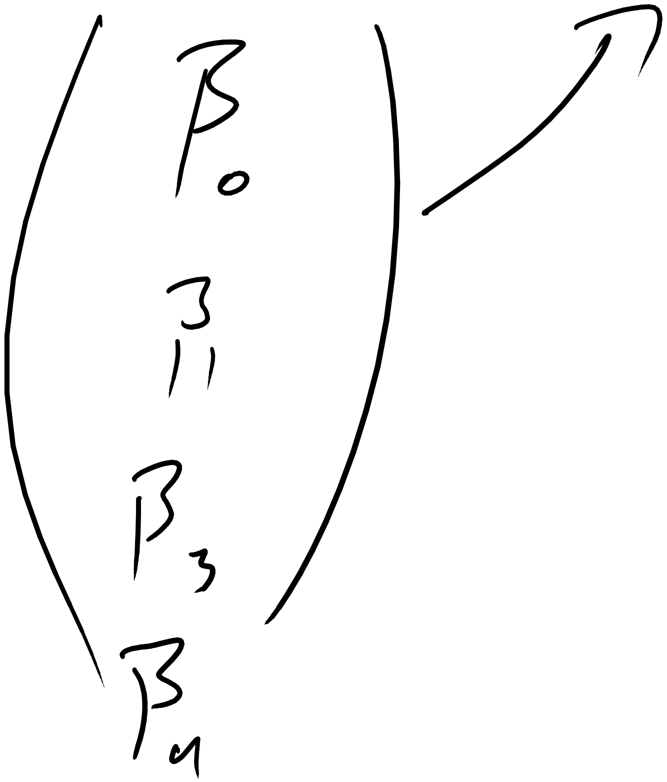
Linear Regression.

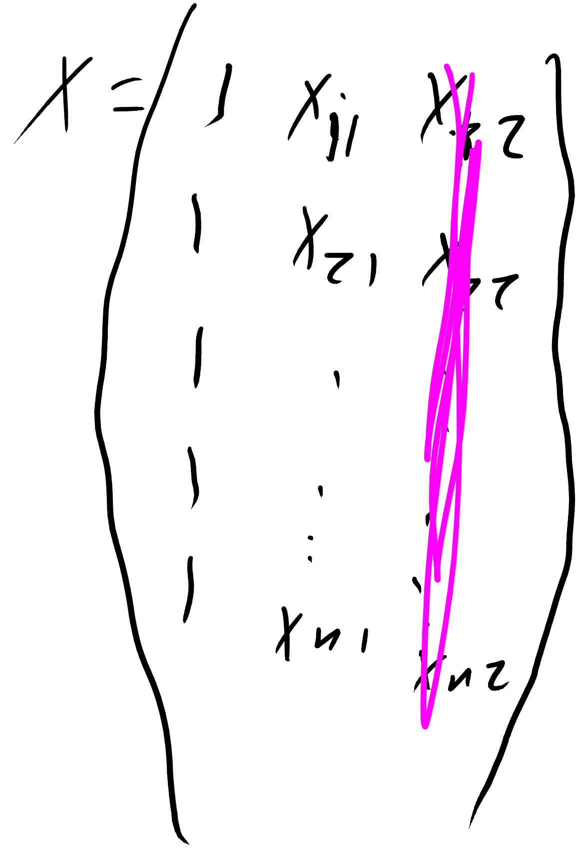


Pop. β_0 β_1
$(x_i, y_i)_{i=1}^n$
LR → $\hat{\beta}_0$ $\hat{\beta}_1$
(x_i, y_i)
LR → $\hat{\beta}_0$ $\hat{\beta}_1$
(x_i, y_i) → LR → $\hat{\beta}_0$ $\hat{\beta}_1$
$\hat{\beta}_1$ $\hat{\beta}_0$ SE($\hat{\beta}_0$) SE($\hat{\beta}_1$)

Standard Errors Matrix Form

$$Var(\hat{\boldsymbol{\beta}}) = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\hat{\sigma}^2$$





Generalized Linear Models

Large Sample Theory

Let $X_1, \ldots, X_n$ be a random sample from a distribution with parameter θ. Let $\hat{\theta}$ be the MLE estimator for a parameter θ. As $n \to \infty$, then $\hat{\theta}$ has a normal distribution with mean θ and variance $1/nI(\theta)$, where

[Math Processing Error]

$$I(\beta_0) = -E\left(\frac{d^2 \ell(\beta_1)}{d\beta_1^2}\right)$$

$$\sqrt{1/nI(\beta_0)} = SE(\beta_0) \quad \text{GLM}$$

Sampling Distributions

Sampling Distributions GLM

ϕ known Binomial/Poisson

$$\frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim N(0,1)$$

$$\hat{\beta}_i \sim N(\beta_i, SE(\hat{\beta}_i)^2)$$

ϕ unknown Normal/Gamma

$$\frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim t_{n-p'}$$

$p' = p + 1$

p = # predictors

$X_1, X_2, X_3 \ldots$