Quant III

Lab 3: GLM and Simulation

Junlong Aaron Zhou

September 25, 2020

Outline

- LR Test
- Marginal Effect
- Monte Carlo and Bootstrap

Likelihood Ratio Tests

- Likelihood: $L(\theta; X) = \prod f_\theta(x_i)$
- Recall what MLE estimator tells us: for which value of parameter can our data be most likely.
- We also want to test some hypothesis: $H_0 : \theta \in \Theta_0$
- LR statistics: $\lambda(y) = \frac{sup_{\theta \in \Theta_0} L(\theta; y)}{sup_{\theta \in \Theta} L(\theta; y)}$, where $\Theta_0 \subset \Theta$
- $-2ln(\lambda(y)) \sim \chi^2_{number\ of\ restriction}$

Likelihood Ratio Tests

- Likelihood: $L(\theta; X) = \prod f_\theta(x_i)$
- Recall what MLE estimator tells us: for which value of parameter can our data be most likely.
- We also want to test some hypothesis: $H_0 : \theta \in \Theta_0$
- LR statistics: $\lambda(y) = \frac{sup_{\theta \in \Theta_0} L(\theta; y)}{sup_{\theta \in \Theta} L(\theta; y)}$, where $\Theta_0 \subset \Theta$
- $-2ln(\lambda(y)) \sim \chi^2_{number\ \ of\ \ restriction}$
- Note: work for nested parameter set: $\Theta_0 \subset \Theta$
- Other tests: Wald statistics, Lagrange multiplier test, etc.

Marginal Effect: Linear Regression

- $Y_i = \alpha + \beta X_i + \epsilon$
- We know $\beta = E(Y_i|X_i = x+1) - E(Y_i|X_i = x)$
- Also, $\beta = \frac{\partial E(Y_i)}{\partial X_i}$
- Marginal Effect.

GLM

- Consider Poisson case:
- $Y_i \sim Poisson(\lambda_i)$
- $\lambda_i = g^{-1}(X_i\beta)$
- What is the link function here?

GLM ctd.

- Poisson has a log link function
- $X_i\beta = \eta_i = log(\lambda_i)$
- Therefore, $\lambda_i = e^{X_i\beta}$
- In most of cases, we are also modeling $E(Y|X) = \lambda_i$
- Eg. Binomial, Normal, etc.
- What is marginal effect of X_i now?

Marginal Effects

$$
\begin{aligned}
ME &= \frac{\partial E(Y_i)}{\partial X_i} \\
&= \frac{\partial E(Y_i)}{\partial \lambda_i} \frac{\partial \lambda_i}{\partial X_i} \\
&= \frac{\partial \lambda_i}{\partial \lambda_i} \frac{\partial e^{X_i \beta}}{\partial X_i} \\
&= \beta e^{X_i \beta}
\end{aligned}
$$

Effects for averages vs average effects

- ME's at sample means: covariates set at sample means.

$$\frac{\partial}{\partial x} \Pr(Y = 1|\bar{x}, \bar{z}_i)$$

- Sample averaged ME's (SAME): use sample values of the covariates and average them.

$$SAME(x) = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial x} \Pr(Y = 1|x, z_i)$$

- Let q be a quantity of interest:

$$q(\hat{\beta}, \bar{x}) \neq \frac{1}{n} \sum_{x} q(\hat{\beta}, x)$$

Interaction term

- What if we have $\lambda = g^{-1}(\beta_1 x1 + \beta_2 x_2 + \beta_3 x_1 * x_2)$

$$
\begin{aligned}
ME &= \frac{\partial E(Y_i)}{\partial X_i} \\
&= \frac{\partial e^{\beta_1 x1 + \beta_2 x_2 + \beta_3 x_1 * x_2}}{\partial x_i} \\
&= (\beta_1 + \beta_3 x_2) e^{\beta_1 x1 + \beta_2 x_2 + \beta_3 x_1 * x_2}
\end{aligned}
$$

Monte Carlo Simulation

- Object: Estimate some parameter of interests of a given distribution.
- For example: expectation, quantile (5%, 95%)
- Notice: $E(X) = \int xf(x)dx$
- Hard to compute when $f(x)$ is ugly.
- $E(g(x)) = \int g(x)f(x)dx$ and $g(x)$ is complex.

Monte Carlo Simulation ctd.

- Idea: if we can sample from $f(x)$, we can use sample mean to approximate population expectation.
- Monte-Carlo Simulation
- More about sampling when we discuss MCMC.

Bootstrap

- Goal: to estimate uncertainty associated with an estimator.
- MLE for point estimates and asymptotic variance via Fisher information
- What if Hessian is too hard to compute? Or what if we have non-parametric regression?
- There is a common method applicable to almost any estimators: **bootstrap**.
- Bootstrap is a data-driven simulation method for statistical inference

Bootstrap: Motivation

- Where does it come from?
- $X \sim F(X|\theta) \rightarrow \{x_i\}_{i=1}^n$
- $\hat{\theta} = h(x_1, \ldots, x_n)$
- Different samples from $F(X|\theta)$ produce different $\hat{\theta}$s that estimate the true θ
- Ideally: take different $\{x_i\}_{i=1}^n$ from $F(X|\theta)$, compute $\hat{\theta}$ for each of them, and get $\hat{\sigma}^2_{\hat{\theta}}$
- But we often have just one sample...
- So, we simulate samples! Parametrically or non-parametrically

Bootstrap: nonparametric

- Instead of $X \sim F(X|\theta)$, assume $X \sim \hat{F}(X|\theta)$
- Our previous sample becomes the new population from which we sample
- Algorithm:
  - Choose B, number of pseudo-samples
  - Sample $\{x_1^{(1)}, \ldots, x_n^{(1)}\}, \ldots, \{x_1^{(B)}, \ldots, x_n^{(B)}\}$
  - Compute $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(B)}$
- $\hat{\sigma}^{*2} = \frac{1}{B-1} \sum_{j=1}^{B} (\hat{\theta}^{(j)} - \bar{\hat{\theta}})^2$, where $\bar{\hat{\theta}} = \frac{1}{B} \sum_{j=1}^{B} \hat{\theta}^{(j)}$
- $(1-\alpha)\%$ CI: cut off $\frac{\alpha}{2}\%$ smallest and largest $\hat{\theta}^{(j)}$ values

Bootstrap: parametric

- Plug $\hat{\theta}$ into $F(X|\theta)$
- Simulate $X \sim F(X|\hat{\theta})$
- Algorithm:
  - Choose B, number of pseudo-samples
  - Sample $\{x_1^{(1)}, \ldots, x_n^{(1)}\}, \ldots, \{x_1^{(B)}, \ldots, x_n^{(B)}\}$ from $F(X|\hat{\theta})$
  - Compute $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(B)}$
- $\hat{\sigma}^{*2} = \frac{1}{B-1} \sum_{j=1}^{B} (\hat{\theta}^{(j)} - \bar{\hat{\theta}})^2$, where $\bar{\hat{\theta}} = \frac{1}{B} \sum_{j=1}^{B} \hat{\theta}^{(j)}$

Difference

- Suppose we know the distribution:
  - Sample directly from the distribution: Monte Carlo
  - By law of large number
  - Sample statistics $\xrightarrow{p}$ Population statistics

Difference

- Suppose we know the distribution:
  - Sample directly from the distribution: Monte Carlo
  - By law of large number
  - Sample statistics $\xrightarrow{p}$ Population statistics
- Suppose we don't know the distribution but empirical distribution
  - Approximate distribution with empirical distribution: bootstrap
  - Either directly sample from empirical distribution (nonparametric), or sample from distribution given estimated parameter (parametric).