Topics4

Saturday, February 27, 2021 6:18 PM

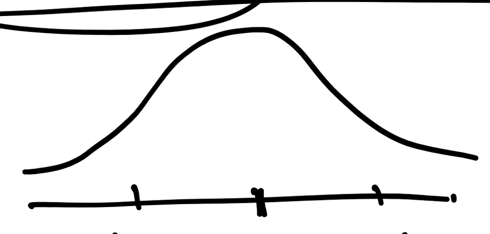
Data → Uni toss | Model: Benulli. | Visualizate.

Data (observed) Model (Theory)

$X_1, X_2, \dots X_n \sim \begin{cases} f(x) & \text{cont.} \\ P(x) & \text{disc.} \end{cases}$

Ex: $X_1, X_2 \dots X_n \sim \text{Normal}(\mu, \sigma^2)$

Data:



-1σ μ $+1\sigma$

average

σ: std. dev. / std. er.

σ^2: variance

Stand normal: $N(0,1)$

Sample mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, n: total number of obs.

Asymptotic Statistics (what's going on when n is large)

MLE →

max. likeli. estimat.

How to find MLE?

Likelihood function: $L(x;\theta) = \dots$

$\frac{\partial}{\partial\theta} \log(L(x;\theta)) \underset{\text{set}}{=} 0 \Rightarrow$ solve for $\hat{\theta}$.

θ: is the parameter

What is the dist. of $(\bar{X})$? (when $n \to \infty$)

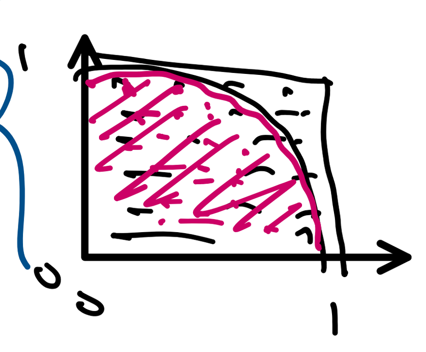
First. we compute mean of $\bar{X}$.

remark ✓ mean ✓ $E(\cdot)$ ✓ 1st moment

$\Rightarrow \begin{cases} E\bar{X} = ? \leftarrow \text{HW} \\ = \mu \leftarrow \text{theoretical mean} \end{cases}$

eq. (1)

ch. π ≈ theorem π



Next. we want var of $\bar{X}$.

remark: variance $\text{var}(\cdot)$ 2nd moment

$\Rightarrow \begin{cases} \text{var}(\bar{X}) = ? \leftarrow \text{HW} \\ = \frac{\sigma^2}{n} \leftarrow \text{a function of theoretical var.} \end{cases}$

eq. (2)

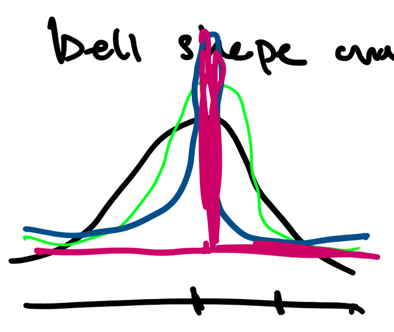
Conclusion (inference):

Central Limit Thm:

$\bar{X} - \mu \rightarrow N(0, \frac{\sigma^2}{n})$

eq. (1) eq. (2)

bell shape curve



μ σ

$0 \approx \frac{\sigma^2 = 1}{n = 10,000}$ $\frac{\sigma^2}{n} \to 0$. if $n \to \infty$.

↑ degenerating variance!!

Update: $\bar{X} - \mu \rightarrow \frac{1}{\sqrt{n}} N(0, \sigma^2)$

" $\sqrt{\ }$ is because PDF of normal dist "

$\Rightarrow \boxed{\sqrt{n}(\bar{X} - \mu) \rightarrow N(0, \sigma^2)}$

asymptotic dist of the estimator $\bar{X}$

In ML:

obs: X : independent var.

Y : dep. var.

Is it possible: $Y \sim f(X) = \hat{Y}$ (target) educated guess $\hat{Y} = f(x) = \beta \cdot X$.

$f: X \to Y$ with obj function Loss Funct. (minimized)

Defn: loss function: Mean Sq. Error (MSE).

→ Y: real target.

$\hat{Y} = f(x)$: guess. $f(x)$: a model

i : running index. $i = 1, 2, \dots n$.

n: total number of obs. in data

→ $\boxed{\text{MSE is defined to be } \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$

loss funct: ex: Mean Sq. Err.

remark: Goal: $\min \arg_{\beta} \cdot \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$

recall how to find MLE

I want to minimize the argument of a loss function by changing β.

Gradient Descent:

Since $\hat{Y} = \beta X$. Because loss is MLE

$L(Y, \hat{Y}) = L(Y, \beta X) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$.

$\nabla L(Y, \hat{Y}) = \frac{\partial}{\partial\beta}\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$.

$= \frac{2}{n}\sum_{i=1}^{n}(Y_i - \beta X)(-\beta)$

each step we update β:

"s"

Grad. Desc $\begin{cases} \beta_{s+1} = \beta_s - \eta \nabla L(Y, \hat{Y}) \end{cases}$

learning rate.