## Review 3

### Estimating probabilities

#### Markov's inequality.

X - non-negative. $P(X \geq a) \leq \frac{\mathbb{E}X}{a}$, $a > 0$

#### Chebyshev inequality

X - has finite variance $Var(X) < \infty$

$$P(|X - \mathbb{E}X| \geq b) \leq \frac{Var(X)}{b^2}$$

just Markov, $Y = (X - \mathbb{E}X)^2$

### Lower bounds.

$P(X < a) \geq 1 - \frac{\mathbb{E}X}{a}$, $P(|X - \mathbb{E}X| < b) = 1 - \frac{Var(X)}{b^2}$.

---

### Sampling

$X_1, X_2, \dots, X_n$ iid $\mathbb{E}X = \mu$, $Var(X) = \sigma^2$

$\bar{X} = 1/n \sum_{i=1}^{n} X_i$ - sample mean.

$\mathbb{E}\bar{X} = \mu \quad , \quad Var(\bar{X}) = \frac{\sigma^2}{n}$

Chebyshev gives

$$P\left(|\bar{X} - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0$$

<u>Weak Law of Large numbers</u> WWLN

$\forall \varepsilon > 0$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) \xrightarrow[n\to\infty]{} 0$$

in otherwords $\bar{X} \xrightarrow{P} \mu$

converges in probability.

<u>Central Limit theorem</u> CLT

standardize

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$$

$$F_Z(x) \longrightarrow F_{\mathcal{Z}}(x) \quad \forall x.$$

$N(0,1)$

In otherwords $Z \xrightarrow{d} N(0,1)$

in distribution.

---

## Convergence of RVs

$$X_n \xrightarrow[n\to\infty]{} X$$

Probability

Distribution.

$$P(|X_n - X| \geq \varepsilon) \xrightarrow[n\to\infty]{} 0 \implies F_{X_n}(x) \xrightarrow[n\to\infty]{} F_X(x).$$

stronger

---

## Continuous Mapping Theorem

$g : \mathbb{R} \to \mathbb{R}$ continuous.

$$X_n \xrightarrow{P} X \quad \Downarrow \quad g(X_n) \xrightarrow{P} g(X)$$

$$X_n \xrightarrow{d} X \quad \Downarrow \quad g(X_n) \xrightarrow{d} g(X).$$

---

## Estimation

- You have distribution w/ unknown parameter $\theta$.
- Probe the distribution by sampling.

$$X_1, X_2, \ldots, X_n \quad \text{iid}.$$

<u>Estimator</u>

$$\hat{\Theta} = \hat{\Theta}(X_1, X_2, \ldots, X_n) \quad - \text{RV}.$$

↑ depends on the sample.

- Approximates $\theta$.

<u>Bias</u>

$$B(\hat{\Theta}) = \mathbb{E}\{\hat{\Theta}\} - \theta$$

<u>Ex</u> ① $\hat{\Theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ – unbiased, $\theta = \mu$.

② $\hat{\Theta} = \bar{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ – biased, $\theta = \sigma^2$

③ $\hat{\Theta} = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ – unbiased, $\theta = \sigma^2$

↑ Bessels correction.

<u>Mean-Square-Error (MSE)</u>

$$MSE(\hat{\Theta}) = \mathbb{E}\{(\hat{\Theta} - \Theta)^2\}$$

$$= Var(\hat{\Theta}) + (B(\hat{\Theta}))^2.$$

Efficiency $\hat{\Theta}_1$ is considered "better" than $\hat{\Theta}_2$
more efficient

if

$$MSE(\hat{\Theta}_1) < MSE(\hat{\Theta}_2).$$

---

Consistency. $\hat{\Theta}_n = \hat{\Theta}(X_1, X_2, \dots, X_n)$.

is consistent if.

$$\hat{\Theta}_n \xrightarrow[n\to\infty]{P} \Theta.$$

Ex

① $\bar{X}$ is consistent (for $\mu$) by WLLN.

② $$\bar{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2$$

Assume $Var(X_i^2) < \infty$ by WLLN

$$1/n \sum_{i=1}^{n} X_i^2 \xrightarrow[n\to\infty]{P} \mathbb{E}X^2, \quad 1/n \sum_{i=1}^{n} X_i \xrightarrow{P} \mathbb{E}X$$

Therefore

$$\bar{S}^2 \xrightarrow{P} \mathbb{E}X^2 - (\mathbb{E}X)^2 = Var(X) = \sigma^2.$$

Note: $S^2 = \underbrace{\frac{n-1}{n}}_{\substack{\downarrow n\to\infty \\ 1}} \bar{S}^2 \Rightarrow S^2$ also consistent.

---

## Maximum Likelihood Estimators (MLE).

① $X_1, X_2, \ldots, X_n$ iid from $f_X(x;\theta)$

unknown ..

② Likelihood function

$$L(\theta) = L(x_1, x_2, \ldots, x_n; \theta) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

③ By iid,

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

$$= f_X(x_1;\theta) f_X(x_2;\theta) \cdots f(x_n;\theta)$$

④ Find $\theta$ which maximizes $L(\theta)$.

$$\theta_{MLE} = h(x_1, x_2, \ldots, x_n)$$

$$\hat{\theta}_{MLE} = h(X_1, X_2, \ldots, X_n).$$

Typically take the log of $L(\theta)$.

$$\ln(AB) = \ln(A) + \ln(B).$$

Log-likelihood.

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

- any maximizer for $\ln(L(\theta))$ is also a maximizer for $L(\theta)$.

---

Ex $Y \sim \text{Geometric}(\theta)$, $P_Y(y;\theta) = \theta(1-\theta)^{y-1}$

$y = 1, 2, 3, \ldots$

MLE for $\theta$?

$$L(\theta) = \prod_{i=1}^{n} \theta(1-\theta)^{y_i-1} = \theta^n \prod_{i=1}^{n}(1-\theta)^{y_i-1}$$

$$\ln(L(\theta)) = \ln\left(\theta^n \prod_{i=1}^{n}(1-\theta)^{y_i-1}\right)$$

$\ln(a^b) = b\ln(a)$

$$= n\ln(\theta) + \sum_{i=1}^{n} \ln(1-\theta)(y_i-1)$$

$$= n\ln\theta + \ln(1-\theta)\sum_{i=1}^{n}(y_i-1)$$

Maximize by $\frac{d}{d\theta} = 0$,

$$\frac{d}{d\theta} \ln(L(\theta)) = \frac{n}{\theta} - \frac{\sum_{i=1}^{n}(y_i - 1)}{1-\theta} = 0$$

$$\Downarrow$$

$$1-\theta - \theta\left(\frac{1}{n}\sum_{i=1}^{n}(y_i - 1)\right) = 0$$

$$1 - \theta\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) = 0$$

$$\theta_{MLE} = \frac{1}{\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)}.$$

" $\mathbb{E}Y = 1/p$.

$p = 1/\mathbb{E}Y$ "

---

## MLE Properties

① Consistent. , ② $B(\hat{\theta}_{MLE}) \xrightarrow{P} 0$. ③ CLT.

$$\frac{\hat{\theta}_{MLE} - \mathbb{E}\hat{\theta}_{MLE}}{\sqrt{Var(\hat{\theta}_{MLE})}} \xrightarrow{d} N(0,1).$$

---

# Confidence intervals.

**Def** $1-\alpha$ % confidence interval $[\hat{\Theta}_L, \hat{\Theta}_u]$.

$$P(\hat{\Theta}_L \le \theta \le \hat{\Theta}_u) \ge 1-\alpha.$$

**Method.**

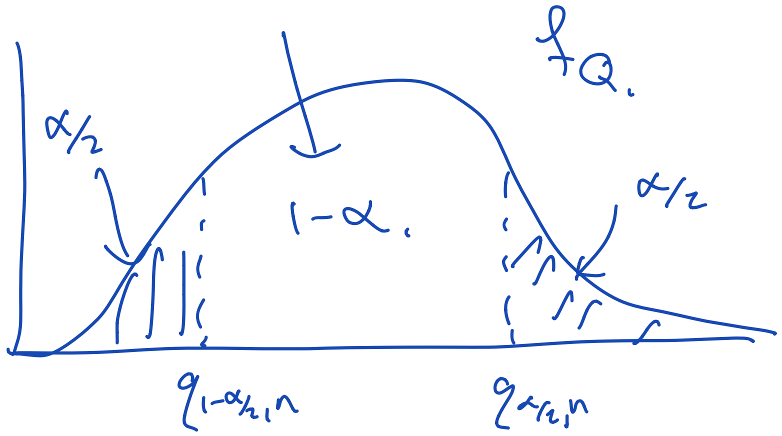
① Pivotal Quantities

$$Q = G(X_1, \dots, X_n, \theta)$$

so that the distribution doesn't depend on $\theta$!

② $$P(q_{\alpha/2,n} \le Q \le q_{1-\alpha/2,n}) = 1-\alpha.$$

$$q_{\alpha/2,n} = F_Q^{-1}(1-\alpha/2)$$




③ Use algebra to find $\hat{\Theta}_L$, $\hat{\Theta}_u$

Ex ① The mean $\mu$.

$$Q = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$ Pivotal quantity.

Could also take $\bar{X} - \mu$.

$Q$ allows for CLT approach if $n \to \infty$,

For the mean. :

If $n$ is big enough.

might be unknown.

$$\left[ \bar{X} - \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}},\ \bar{X} + \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}} \right]$$ $1-\alpha\,\%$ confidence interval for $\mu$.

If $\sigma$ is unknown

① can bound by $\sigma \leq \sigma_{max}$.

② can be replaced by $S^2$

② implies.

$$\left[\bar{X} - \frac{z_{\alpha/2} S}{\sqrt{n}},\ \bar{X} + \frac{z_{\alpha/2} S}{\sqrt{n}}\right].$$

is a $1-\alpha$ % confidence interval.

---

## Estimating variance for normal distributions,

$$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$$

Pivotal quantity.

$$Q = \frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1)$$

$$P\left(\chi^2_{\alpha/2, n-1} \le Q \le \chi^2_{1-\alpha/2, n-1}\right) = 1-\alpha.$$

↳ Algebra,

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}},\ \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right]$$ $1-\alpha$ confidence interval.

$\chi^2_{\alpha/2, n-1}$ = chi2inv$(1-\alpha/2, n-1)$.

or use table on webpage

Estimating the mean when $n$ is not big.

$X_1, X_2, \ldots, X_n \sim N(0,1)$.

normal sample.

① $Q = \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim T(n-1)$

t-distribution,

② $P\left(t_{\alpha/2, n-1} \leq Q \leq t_{1-\alpha/2, n-1}\right) = 1 - \alpha$

$t_{\alpha/2, n-1} = \text{tinv}(1 - \alpha/2, n-1)$

look up in table

③ Algebra.

$$\left[\bar{X} - \frac{t_{\alpha/2, n-1} S}{\sqrt{n}}, \; \bar{X} + \frac{t_{\alpha/2, n-1} S}{\sqrt{n}}\right].$$

$1-\alpha$ % confidence interval.