Stat 135 lec 4

Last time

We showed the highlighted part of the table below.

usually we don't know σ^2 so we use $(1-\frac{1}{N})s^2$ instead

Population Parameter	Estimate	Variance of Estimate	Estimated Variance
μ	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$	$s_{\overline{X}}^2 = \frac{s^2}{n}\left(1-\frac{n}{N}\right)$
p	$\hat{p}$ = sample proportion	$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)$	$s_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n-1}\left(1-\frac{n}{N}\right)$
τ	$T = N\overline{X}$	$\sigma_T^2 = N^2\sigma_{\overline{X}}^2$	$s_T^2 = N^2 s_{\overline{X}}^2$
σ^2	$\left(1-\frac{1}{N}\right)s^2$		

(2) We motivated why we estimate parameters of a probability model.

P256 Rice

data

Pois($\lambda = 8.4$) prob model for # intervals with n α particle arrivals.

n	Observed	Expected
0–2	18	12.2
3	28	27.0
4	56	56.5
5	105	94.9
6	126	132.7
7	146	159.1
8	164	166.9
9	161	155.6
10	123	130.6
11	101	99.7
12	74	69.7
13	53	45.0
14	23	27.0
15	15	15.1
16	9	7.9
17+	5	7.1
	1207	1207

Today sec 8.4 Method of moment (MOM) estimators

Consistency of MOM estimators

Sec 8.4 The Method of Moments (MOM)

Review of Gamma Distribution

$X \sim \text{Gamma}(r, \lambda)$ — r: rth particle; λ: rate of arrival of a particle in Pois(λ) process.

X: time to the r^{th} arrival of a Pois(λ) process.

$$f(x) = \frac{\lambda^r}{\Gamma(r)} \underbrace{x^{r-1} e^{-\lambda x}}_{\text{variable part of } f(x)} \qquad \Gamma(r) = (r-1)! \text{ for } r \text{ a pos integer}$$

$$E(X) = \frac{r}{\lambda}$$

$$\text{Var}(X) = r/\lambda^2$$

MOM

If we want to estimate ℓ parameters $(\theta_1, \ldots, \theta_\ell)$ of a prob dist $f(x \mid \theta_1, \ldots, \theta_\ell)$ from i.i.d sample $X_1, \ldots, X_n$ from this distribution

Step 1 Compute the first ℓ moments

$$\mu_k = E(X^k) \qquad k = 1, \ldots, \ell$$

$$\mu_k = \int_{-\infty}^{\infty} x^k f(x \mid \theta_1, \ldots, \theta_\ell)\, dx$$

$$\mu_1 = g_1(\theta_1, \ldots, \theta_\ell)$$

$$\mu_2 = g_2(\theta_1, \ldots, \theta_\ell)$$

$$\vdots$$

$$\mu_\ell = g_\ell(\theta_1, \ldots, \theta_\ell)$$

Ex $X \sim \text{Pois}(\lambda)$ $\quad \ell = 1$ $\quad X = \#$ arrivals in 10 sec intervals

$$\mu_1 = E(X) = \lambda$$

Ex $X \sim \text{Gamma}(r, \lambda)$ $\ell = 2$

$\mu_1 = E(X) = r/\lambda$

$$\mu_2 = E(X^2) = \underset{\underset{r/\lambda^2}{\shortparallel}}{\text{var}(X)} + \underset{\underset{\left(\frac{r}{\lambda}\right)^2}{\shortparallel}}{E(X)^2} = \frac{r + r^2}{\lambda^2}$$

Step 2 Use algebra to invert the above system of equations

$$\left.\begin{aligned} \theta_1 &= h_1(\mu_1, \ldots, \mu_\ell) \\ \theta_2 &= h_2(\mu_1, \ldots, \mu_\ell) \\ &\vdots \\ \theta_\ell &= h_\ell(\mu_1, \ldots, \mu_\ell) \end{aligned}\right\} \quad (*)$$

require h to be a continuous function of $\mu_1, \ldots, \mu_\ell$.

Ex Pois case

$\mu_1 = \lambda \Rightarrow \lambda = \mu_1$

Ex Gamma case

$\mu_1 = \frac{r}{\lambda}$

$$\mu_2 = \frac{r}{\lambda^2} + \frac{r^2}{\lambda^2} = \frac{\mu_1}{\lambda} + \mu_1^2$$

$$\Rightarrow \frac{\mu_1}{\lambda} = \mu_2 - \mu_1^2$$

$$\boxed{\lambda = \frac{\mu_1}{\mu_2 - \mu_1^2}}$$

$$\boxed{r = \lambda \mu_1 = \frac{\mu_1^2}{\mu_2 - \mu_1^2}}$$

Step 3

We insert into (*) the estimator for the moments $\mu_1, \dots, \mu_\ell$ (sample moments)

$$\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

You will show in HW that these are unbiased estimators of $\mu_1, \dots, \mu_\ell$

$$\hat{\mu}_2 = \frac{1}{n}\sum_{i}^{n} X_i^2$$

$E(X^2)$

$\vdots$

$$\hat{\mu}_\ell = \frac{1}{n}\sum_{i=1}^{n} X_i^\ell$$

$$\hat{\theta}_1 = h_1(\hat{\mu}_1, \dots, \hat{\mu}_\ell)$$
$$\hat{\theta}_2 = h_2(\hat{\mu}_1, \dots, \hat{\mu}_\ell)$$
$$\vdots$$
$$\hat{\theta}_\ell = h_\ell(\hat{\mu}_1, \dots, \hat{\mu}_\ell)$$

MoM estimator for $\theta_1, \dots, \theta_\ell$

ex Pois $\lambda = \mu_1 \Rightarrow \hat{\lambda} = \hat{\mu}_1 = \bar{X}$

ex Gamma
$$\boxed{\hat{\lambda} = \frac{\hat{\mu}_1}{\hat{\mu}_2 - \hat{\mu}_1^2} \qquad \hat{r} = \frac{\hat{\mu}_1^2}{\hat{\mu}_2 - \hat{\mu}_1^2}}$$

Next we will show MOM estimators are consistent $\hat{\Theta}_{mom} \xrightarrow{P} \Theta$

converge in probability.

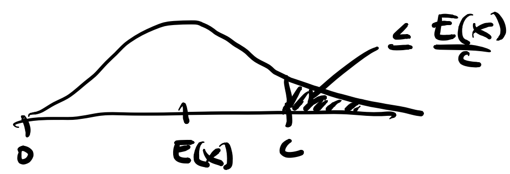
A little Background

Proof of WLLN:

<u>Markov's Inequality</u>

For $X \ge 0$, $c > 0$

$$P(X \ge c) \le \frac{E(X)}{c}$$



eg let $X_1, \dots, X_n$ iid w mean μ, var σ^2

$\bar{X}_{(n)} = \frac{1}{n} \sum_{i=1}^{n} X_i$ <u>note</u> $E(\bar{X}_{(n)}) = \mu$

$Var(\bar{X}_{(n)}) = \frac{\sigma^2}{n}$

Let $\varepsilon > 0$

Use Markov's ineq to give an upper bound for

$$P(\underbrace{|\bar{X}_{(n)} - \mu|}_{\text{a nonneg RV}} \ge \varepsilon)$$

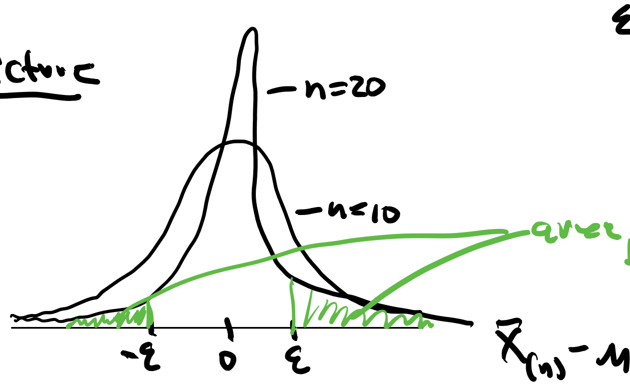
pos constant

$$P(|\bar{X}_{(n)} - \mu| \ge \varepsilon) = P((\bar{X}_{(n)} - \mu)^2 \ge \varepsilon^2)$$

$$\le \frac{E((\bar{X}_{(n)} - \mu)^2)}{\varepsilon^2} = \frac{Var(\bar{X}_{(n)})}{\varepsilon^2}$$

$$= \frac{\sigma^2}{n\varepsilon^2}$$

<u>Picture</u>



area in tails is bound by $\frac{\sigma^2}{n\varepsilon^2} \to 0$ as $n \to \infty$.

<u>Defⁿ</u> (Convergence in Probability) — the probability of an unusual outcome becomes smaller and smaller

A sequence $X_1, X_2, \dots$ of RV converge in prob to a RV X if for any $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \to 0 \quad \text{as } n \to \infty.$$

ex Weak law of large numbers (WLLN)

$X_1, \dots, X_n$ iid w/ mean μ, var σ^2

$\overline{X}_{(n)} \xrightarrow{P} \mu$ ← here μ is the constant RV that takes value μ with prob 1.

This follows directly from Markov's Ineq.

$$P(|\overline{X}_{(n)} - \mu| > \varepsilon) \le \frac{\sigma^2}{\varepsilon^2 \cdot n} \to 0 \quad \text{as } n \to \infty.$$

ex Generalization of WLLN to higher moments

$X_1, \dots, X_n$ iid w/ mean μ, var σ^2

$$\underbrace{\hat{\mu}_k}_{= \frac{1}{n}\sum_{i=1}^{n} X^k} \xrightarrow{P} \underbrace{\mu_k}_{E(X^k)}$$

<u>Defⁿ</u> An estimator $\hat{\theta}$ of a parameter θ is <u>consistent</u> if $\hat{\theta} \xrightarrow{P} \theta$

We argue that MOM is consistent.

<u>Thm</u> Suppose RVs $X_1, X_2, \dots$ converge in prob to a RV X and h is a continuous function, Then $h(X_1), h(X_2), \dots$ converge in prob to $h(X)$

So if h is continuous then the estimator is consistent

$$\hat{\theta}_{MoM} = h(\hat{\mu}_1, \dots, \hat{\mu}_\ell) \xrightarrow{P} \theta$$

$$= h(\mu_1, \dots, \mu_\ell)$$