02.11.2021

Seminar W6 — 831

X random variable, $(X_n)_{n\in\mathbb{N}}$ sequence of random variables that are i.i.d. (independent and identically distributed), $X_n \sim X$

Statistic: $h: \mathbb{R}^n \to \mathbb{R}$ measurable

$$h = h(X_1, \dots, X_n)$$

$\underline{\varphi_X}$:
$$\overline{X} = \frac{1}{n}\cdot(X_1 + X_2 + \dots + X_n)$$
$$\overline{\nu_k} = \frac{1}{n}\left(X_1^k + X_2^k + \dots + X_n^k\right)$$
$$\overline{\mu_k} = \frac{1}{n}\left((X_1 - \overline{X})^k + \dots + (X_n - \overline{X})^k\right)$$

Exercise 1. Let $X_1, X_2, \dots, X_n, \dots$ be i.i.d. (independent identically distributed) random variables that follow the normal distribution, $X \sim \mathcal{N}(\mu, \sigma)$. Find the constant k_n such that the sampling function

$$\overline{s} = k_n \sum_{j=1}^{n} |X_j - \overline{X}|$$

verifies $E(\overline{s}) = \sigma$.

Sol.: $\forall X, Y$: $E(X+Y) = E(X) + E(Y)$

$\forall \alpha \in \mathbb{R}$: $E(\alpha X) = \alpha E(X)$

$$E(\overline{s}) = E\left(k_n \cdot \sum_{j=1}^{n} |X_j - \overline{X}|\right) = k_n \cdot E\left(\sum_{j=1}^{n} |X_j - \overline{X}|\right)$$

$$\overline{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

$$E(\overline{S}) = k_n \cdot \sum_{j=1}^{n} E\left(|X_j - \overline{X}|\right)$$

$X_1, \dots, X_n$ identically distributed $\Rightarrow E\left(|X_j - \overline{X}|\right) = E(|X_i - \overline{X}|)$ $\forall i, j$

$$E(\overline{S}) = k_n \cdot n \cdot E\left(|X_1 - \overline{X}|\right)$$

$X_1 \sim \mathcal{N}(\mu, \sigma)$, $E(X_1) = \mu$
$\sigma_{X_1} = \sqrt{V(X_1)} = \sigma$

Let $Y = X_1 - \overline{X}$

$$X_1 \sim \mathcal{N}(\mu, \sigma)$$

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

If X and Y are independent random variables with $X, Y \sim \mathcal{N}(\mu, \sigma)$, then for any $\alpha, \beta \in \mathbb{R}$:

$$\alpha X + \beta Y \sim \mathcal{N}(\mu', \sigma')$$

$$Y = X_1 - \overline{X} = X_1 - \frac{1}{n} \cdot (X_1 + X_2 + \dots + X_n) =$$

$$= \frac{n-1}{n} \cdot X_1 + \frac{1}{n} X_2 + \dots + \frac{1}{n} X_n$$

$$\Rightarrow Y \sim \mathcal{N}(\mu', \sigma')$$

$$\mu' = E(Y) = E(X_1 - \overline{X}) = E(X_1) - E(\overline{X}) = \mu - \mu = 0$$

$$E(\bar{X}) = E\left(\frac{1}{n}(X_1 + \dots + X_n)\right) = \frac{1}{n}\cdot\left(E(X_1) + E(X_2) + \dots + E(X_n)\right) =$$

$$= \frac{1}{n}\cdot n \cdot \underbrace{E(X_1)}_{=\mu} = \mu$$

$$\sigma' = \sqrt{V(Y)}$$

$$V(Y) = V(X_1 - \bar{X}) = V\left(X_1 - \frac{1}{n}(X_1 + X_2 + \dots + X_n)\right) =$$

$$= V\left(\frac{n-1}{n}X_1 + \left(-\frac{1}{n}\right)X_2 + \dots + \left(-\frac{1}{n}\right)X_n\right) \overset{X_1, \dots, X_n \text{ independent}}{=}$$

$$= V\left(\frac{n-1}{n}X_1\right) + V\left(-\frac{1}{n}X_2\right) + \dots + V\left(-\frac{1}{n}X_n\right) =$$

$$= \left(\frac{n-1}{n}\right)^2 \underbrace{V(X_1)}_{=\sigma^2} + \frac{1}{n^2}\cdot\underbrace{V(X_2)}_{=\sigma^2} + \dots + \frac{1}{n^2}\underbrace{V(X_n)}_{=\sigma^2} =$$

$$= \sigma^2 \cdot \left(\left(\frac{n-1}{n}\right)^2 + (n-1)\cdot\frac{1}{n^2}\right) = \sigma^2 \cdot \frac{n-1}{n^2}(n-1+1) = \sigma^2 \cdot \frac{n-1}{n}$$

$$\Rightarrow Y \sim \mathcal{N}(\mu', \sigma') \quad \text{and} \quad \mu' = 0 \qquad \sigma' = \sigma\sqrt{\frac{n-1}{n}}$$

$$\Rightarrow Y \sim \mathcal{N}\left(0, \sigma\sqrt{\frac{n-1}{n}}\right)$$

$$E(\bar{S}) = k_n \cdot n \cdot E(|Y|)$$

$$E(|Y|) = \int_{\mathbb{R}} \underbrace{|y|}_{\text{"the money you gain by playing the game"}} \cdot \underbrace{f_Y(y)}_{\text{the "probability" of gaining that money}} \, dy$$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R} \qquad X \sim \mathcal{N}(\mu, \sigma)$$

$$Y \sim \mathcal{N}\left(0, \sigma\sqrt{\tfrac{n-1}{n}}\right) \Rightarrow f_Y(x) = \frac{1}{\sigma\sqrt{\frac{n-1}{n}} \cdot \sqrt{2\pi}} \cdot e^{-\frac{x^2}{2\sigma^2 \cdot \frac{n-1}{n}}}$$

$$E(|Y|) = \int_{\mathbb{R}} |y| \cdot f_Y(y)\, dy = \underbrace{\int_{-\infty}^{0} (-y) \cdot f_Y(y)\, dy}_{y=-z} +$$

$$+ \int_0^\infty y\, f_Y(y)\, dy$$

$$\int_\infty^0 z \cdot f_Y(-z) \cdot (-dz) =$$

$$= \int_0^\infty z \underbrace{f_Y(-z)}_{f_Y(z)} \cdot dz =$$

$$= \int_0^\infty y\, f_Y(y)\, dy$$

$$= 2\int_0^\infty y\, f_Y(y)\, dy = 2\int_0^\infty y \cdot \frac{1}{\sigma\sqrt{\frac{n-1}{n}} \cdot \sqrt{2\pi}} \cdot e^{\frac{-y^2}{2\sigma^2 \cdot \frac{n-1}{n}}}\, dy =$$

$$= \frac{2}{\sigma\sqrt{\frac{n-1}{n}} \cdot \sqrt{2\pi}} \cdot \int_0^\infty \underbrace{y \cdot e^{\frac{-y^2}{2\sigma^2 \cdot \frac{n-1}{n}}}}\, dy =$$

$$\left(e^{\frac{-y^2}{2\sigma^2 \cdot \frac{n-1}{n}}}\right)' = e^{\frac{-y^2}{2\sigma^2 \cdot \frac{n-1}{n}}} \cdot \frac{-2y}{2\sigma^2 \cdot \frac{n-1}{n}}$$

$$= \frac{2}{\sigma\sqrt{\frac{n-1}{n}} \cdot \sqrt{2\pi}} \cdot \underbrace{\left(e^{\frac{-y^2}{2\sigma^2 \cdot \frac{n-1}{n}}}\right)\Bigg|_0^\infty}_{=-1} \cdot \frac{2\sigma^2 \cdot \frac{n-1}{n}}{-2} =$$

$$= \frac{2\sigma^2 \cdot \frac{n-1}{n}}{\sigma\sqrt{\frac{n-1}{n}} \cdot \sqrt{2\pi}} = \sigma \cdot \sqrt{\frac{n-1}{n}} \cdot \sqrt{\frac{2}{\pi}}$$

$$E(S) = \sigma\sqrt{\frac{n-1}{n}} \cdot \sqrt{\frac{2}{\pi}} \cdot k_n \cdot n = \sigma$$

$$\Rightarrow k_n = \frac{1}{n\sqrt{\frac{n-1}{n}} \cdot \sqrt{\frac{2}{\pi}}}$$

Ex 2 :

Ex. 3 :

Exercise 3. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. (independent identically distributed) random variables that follow the distribution

$$X \sim Unif[a,b]$$

where $0 < a < b$, and consider the following statistics:

1. Arithmetic mean of selection:

$$a_n(X_1,\dots,X_n) := \frac{1}{n}\sum_{i=1}^{n} X_i$$

2. Geometric mean of selection:

$$g_n(X_1,\dots,X_n) := \sqrt[n]{\prod_{i=1}^{n} X_i}$$

3. Harmonic mean of selection:

$$h_n(X_1,\dots,X_n) := \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}$$

Prove that each of the above statistics converges almost surely to a constant, as $n \to \infty$ and find these constants.

Recap.

- *A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables converges* ***almost surely*** *(denoted by* **a.s.** *and written as $X_n \overset{a.s}{\to} X$) to a random variable X if:*

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

- ***The Strong Law of Large Numbers (SLLN):***

  If $(X_n)_{n\in\mathbb{N}}$ is a sequence of i.i.d. random variables with $X_n \sim X$, then

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s} E(X)$$

- *If $X \sim Unif[a,b]$, then:*

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } x \in [a,b] \\ 0, & \text{otherwise} \end{cases}$$

Sol.:

$$a_n(X_1, X_2, \dots, X_n) = \frac{1}{n}(X_1 + X_2 + \dots + X_n) \to E(X)$$

$$X \sim Unif([a,b]), \quad f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}$$

$$E(X) = \int_{\mathbb{R}} x \cdot f_X(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \cdot \left(\frac{1}{2}x^2\right)\Big|_a^b =$$

$$= \frac{1}{b-a} \cdot \frac{1}{2} \cdot (b^2 - a^2) = \frac{a+b}{2}$$

$$\Rightarrow \quad a_n \xrightarrow{a.s.} \frac{a+b}{2}$$

$$g_n = \sqrt[n]{X_1 X_2 \cdots X_n}$$

$$\ln g_n = \frac{1}{n} \cdot (\ln X_1 + \ln X_2 + \dots + \ln X_n)$$

$$Y_n := \ln X_n \Rightarrow Y_n \sim \ln X$$

By the SLLN: $\frac{1}{n}(Y_1 + \dots Y_n) \xrightarrow{a.s.} E(\ln X)$

$$E(\ln X) = \int_{\mathbb{R}} \ln x \cdot f_X(x)\, dx = \int_a^b \ln x \cdot \frac{1}{b-a}\, dx =$$

$$= \frac{1}{b-a} \cdot \int_a^b \ln x\, dx$$

$$\int_a^b \ln x\, dx = \int_a^b x' \ln x\, dx = x \ln x \Big|_a^b - \int_a^b \underbrace{x \cdot (\ln x)'}_{1}\, dx =$$

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

$$\Rightarrow \ln g_n \xrightarrow{a.s.} \frac{1}{b-a}\left(b \ln b - a \ln a - (b-a)\right)$$

$$g_n \xrightarrow{a.s.} e^{\frac{1}{b-a}(b \ln b - a \ln a - (b-a))}$$

$$h_n = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

$$h_n = \left[\frac{1}{n}\left(\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_n}\right)\right]^{-1}$$

$$\frac{1}{X_n} =: Y_n \qquad \Rightarrow \quad Y_n \sim \frac{1}{X} =: Y$$

$$\frac{Y_1 + \dots + Y_n}{n} \xrightarrow{a.s.} E(Y) = E\left(\frac{1}{X}\right)$$

$$\Rightarrow \quad \frac{1}{h_n} \xrightarrow{a.s.} E\left(\frac{1}{X}\right) \quad \Rightarrow \quad h_n \xrightarrow{a.s.} \frac{1}{E\left(\frac{1}{X}\right)}$$

$$E\left(\frac{1}{X}\right) = \int_{\mathbb{R}} \frac{1}{x} \cdot f_X(x)\,dx = \int_a^b \frac{1}{x} \cdot \frac{1}{b-a}\,dx = \frac{1}{b-a} \cdot \ln x \Big|_a^b =$$

$$= \frac{1}{b-a}\left(\ln b - \ln a\right) = \frac{\ln \frac{b}{a}}{b-a}$$