Various Topics: Tues Mar 16, 2021

- “independent and identically distributed”
- Variance and Covariance Properties
- Chebyshev and Markov’s Inequalities
- Weak Law of Large Numbers
- Review problems – CLT, Normal Distributions

Чебышёв

i.i.d.

If we have multiple random variables that come from exactly the same distribution and are (mutually) independent, we can say that these random variables are **i.i.d**.

"**iid**" means "**<u>i</u>ndependent** and **<u>i</u>dentically <u>d</u>istributed**" e.g.

- $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathrm{f(x)}$
- $X_1, X_2, \ldots, X_n \overset{iid}{\sim} Poisson(4)$
- $X_1, X_2, \ldots, X_n \overset{iid}{\sim} Bernoulli(.2)$

Markov's Inequality

Markov's inequality: for any nonnegative random variable X, and for any $t > 0$,

$$\mathbf{Pr}[X \geq t] \leq \frac{\mathbf{E}[X]}{t}.$$

Proof: say X can take values $x_1 < x_2 < \ldots x_j = t < \ldots x_n$

$$E[X] = \sum_{i=1}^{n} x_i \cdot f(x_i) \geq \sum_{i=j}^{n} x_i \cdot f(x_i) \geq \sum_{i=j}^{n} t \cdot f(x_i)$$

$$= t \sum_{i=j}^{n} f(x_i)$$

Continuous proof is similar to this, but with integrals

Markov's inequality: for any nonnegative random variable X, and for any $t > 0$,

$$\mathbf{Pr}[X \geq t] \leq \frac{\mathbf{E}[X]}{t}.$$

let $t = s \cdot E[X]$:

$$\mathbf{Pr}[X \geq s \cdot \mathbf{E}[X]\,] \leq \frac{1}{s}.$$

Chebyshev's Inequality

$P[Y \geq s \cdot E[Y]] \leq \frac{1}{s}$ let $s = k^2$:

$P[Y \geq k^2 \cdot E[Y]] \leq 1/k^2$ let $Y = (X-\mu)^2$

$P[(X-\mu)^2 \geq k^2 \cdot E[X-\mu]^2] \leq \frac{1}{k^2}$

For any random variable **X**, and scalars {*t, a*} ∈ R with t > 0:

$$\mathbf{Pr}[\,|X-a| \geq t] = \mathbf{Pr}[\,(X-a)^2 \geq t^2]$$

$\rightarrow$ $\boxed{P[|X-\mu| > k\sigma] \leq \frac{1}{k^2}}$ $\longleftarrow$ Chebyshev's Inequality

Chebyshev Example

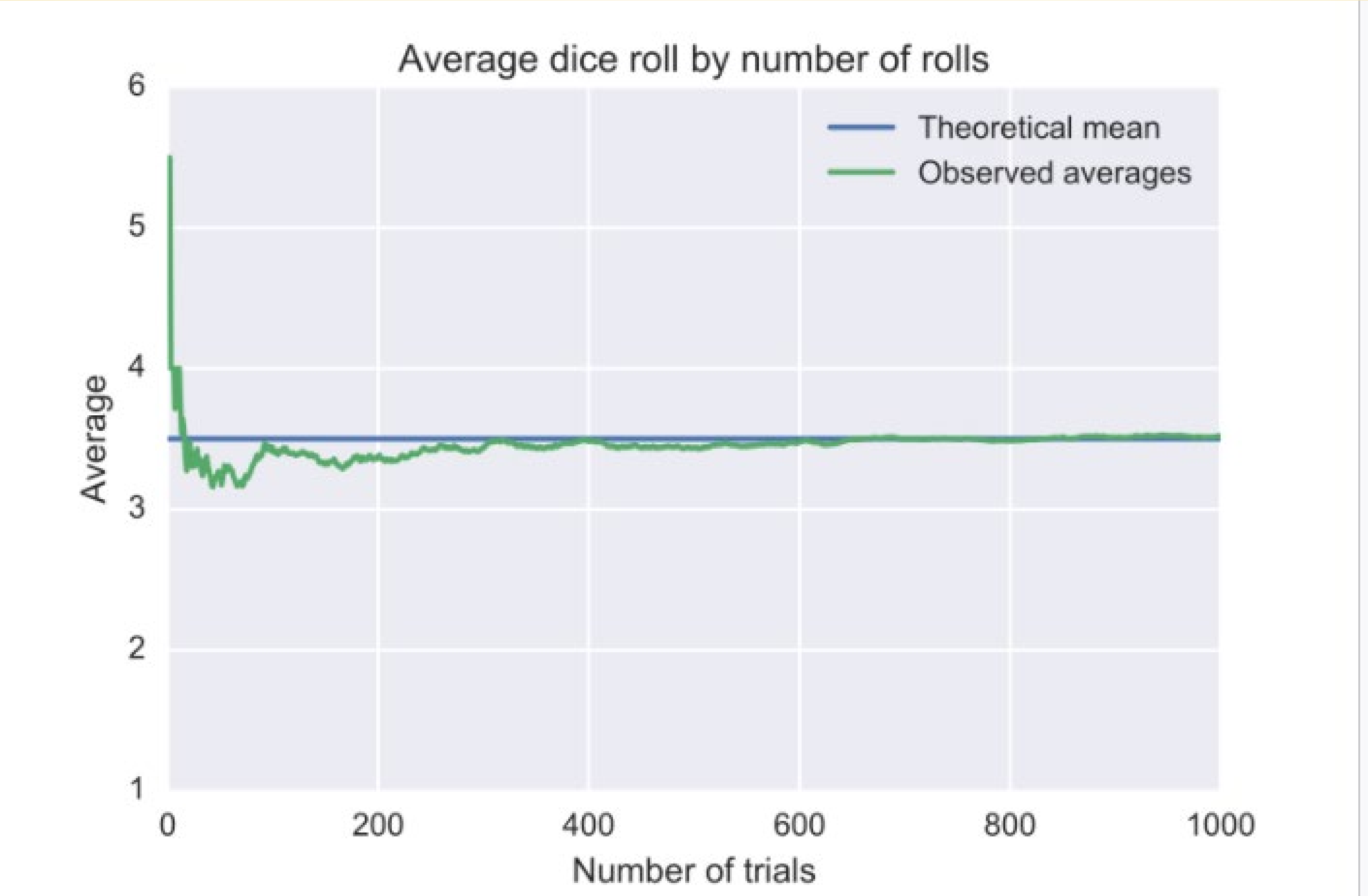
$P\left[|X-\mu| > k\sigma\right] \leq \frac{1}{k^2}$

Find an **upper bound** for the probability that a random variable can be more than 2 standard deviations away from its mean.

notes

Weak Law of Large Numbers (WLLN)

$$\overline{X}_n \xrightarrow{P} \mu \qquad \text{when } n \to \infty.$$



https://en.wikipedia.org/wiki/Law_of_large_numbers

Weak Law of Large Numbers (WLLN)

$$\overline{X}_n \xrightarrow{P} \mu \qquad \text{when } n \to \infty.$$

Let X_1, X_2 , ... , X_n be i.i.d. random variables with a finite expected value $EX_i = \mu < \infty$. Then, for any $\epsilon > 0$,

$$\lim_{n \to \infty} P(|\overline{X} - \mu| \geq \epsilon) = 0.$$

$P[|X-\mu| > k\sigma] \leq \frac{1}{k^2}$ Let $\varepsilon = k\sigma$. Then, $k = \varepsilon/\sigma$ $P[|X-\mu| > \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}$

$$\mathrm{P}(|\overline{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

More Examples (old material)

Variance & Covariance

Functions of Normal Distribution

Central Limit Theorem

Variance Examples (assuming independence)

Let $X_1, X_2, \dots X_n \overset{iid}{\sim} \mathrm{N}(\mu_X, \sigma_X^2)$.

Find in terms of σ_X^2:

1. Var$[X_1 + X_2]$

$= \mathrm{Var}[(1)X_1 + (1)X_2]$

$= (1)^2\mathrm{Var}[X_1] + (1)^2\mathrm{Var}[X_2]$

$= \mathrm{Var}[X_1] + \mathrm{Var}[X_2]$

$= \sigma_X^2 + \sigma_X^2$

$= 2\sigma_X^2$

2. Var$[X_1 + X_2 + X_3]$

$= \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \mathrm{Var}[X_3]$

$= \sigma_X^2 + \sigma_X^2 + \sigma_X^2$

$= 3\sigma_X^2$

Variance Examples

Let $X_1, X_2, \dots X_n \overset{iid}{\sim} \text{N}(\mu_X, \sigma_X^2)$.

Find in terms of σ_X^2:

3. $\text{Var}[X_1 - X_2]$

$= \text{Var}[(1)X_1 + (-1)X_2]$

$= (1)^2\text{Var}[X_1] + (-1)^2\text{Var}[X_2]$

$= \text{Var}[X_1] + \text{Var}[X_2]$

$= \sigma_X^2 + \sigma_X^2 = 2\sigma_X^2$

4. $\text{Var}[X_1 + X_2 + \dots + X_n] = n\sigma_X^2$

5. $\text{Var}[3X_1] = (3)^2\text{Var}[X_1] = 9\sigma_X^2$

6. $\text{Var}[3X_1 - 2X_2 - X_3]$

$= \text{Var}[(3)X_1 + (-2)X_2 + (-1)X_3]$

$= (3)^2\text{Var}[X_1] + (-2)^2\text{Var}[X_2] + (-1)^2\text{Var}[X_3]$

$= 9\sigma_X^2 + 4\sigma_X^2 + 1\sigma_X^2$

$= 14\sigma_X^2$

Variance Examples

Let $X_1, X_2, \dots X_n \overset{iid}{\sim} \mathrm{N}(\mu_X, \sigma_X^2)$.

Find in terms of σ_X^2:

7. $\mathrm{Var}[aX_1 + \mathrm{b}X_5 - \mathrm{c}X_2]$

$= (\mathrm{a})^2\mathrm{Var}[X_1] + (\mathrm{b})^2\mathrm{Var}[X_5] + (\mathrm{-c})^2\mathrm{Var}[X_2]$

$= \mathrm{a}^2\sigma_X^2 \quad + \mathrm{b}^2\sigma_X^2 \quad + \mathrm{c}^2\sigma_X^2$

$= (\mathrm{a}^2 + \mathrm{b}^2 + \mathrm{c}^2)\,\sigma_X^2$

8. $\mathrm{Var}[\bar{X}]$

$= \mathrm{Var}\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right]$

$= \mathrm{Var}\left[\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n\right)\right]$

$= \frac{1}{n^2}\mathrm{Var}[X_1] + \frac{1}{n^2}\mathrm{Var}[X_2] + \ldots + \frac{1}{n^2}\mathrm{Var}[X_2]$

$= \frac{1}{n^2}\underbrace{(\mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \ldots + \mathrm{Var}[X_n])}_{n \text{ terms}}$

$= \frac{1}{n^2}(\mathrm{n}\,\sigma_X^2)$

$= \boxed{\frac{\sigma_X^2}{n}}$ ← important

Covariance Practice

Find in terms of $\sigma_X^2,\ \sigma_Y^2,\ and\ \ \sigma_{XY}$:

1. Var[X+Y] = Cov[X+Y, X+Y]

= Cov[X,X]+Cov[X,Y]+Cov[Y,X]+Cov[Y,Y]
= σ_X^2 + $2\sigma_{XY}$ + σ_Y^2

2. Var[2X -3Y] = Cov[2X + (-3)Y, 2X + (-3)Y]

= Cov[2X, 2X] + Cov[2X, -3Y] + Cov[-3Y, 2X] + Cov[-3Y, -3Y]

= (2)(2)Cov[X,X] + 2(2)(-3)Cov[X,Y] + (-3)(-3)Cov[Y, Y]

= $4\sigma_X^2$ $-12\sigma_{XY}$ $+9\sigma_Y^2$

Covariance Practice

Find in terms of $\sigma_X^2, \sigma_Y^2, \sigma_{XY}$:

3. Cov[3X-Y, 2X + 5Y]

= Cov[3X,2X] + Cov[3X, 5Y] + Cov[-Y, 2X] + Cov[-Y, 5Y]

= $6\sigma_X^2$ + $15\sigma_{XY}$ + $-2\sigma_{XY}$ + $-5\sigma_Y^2$

= $6\sigma_X^2$ + $13\sigma_{XY}$ - $5\sigma_Y^2$

Normal Example

A bartender has Guinness on tap, and is instructed to fill glasses up to an average of μ ounces per glass. The amount that she fills is normally distributed with $\mathbf{mean} = \boldsymbol{\mu}$ ounces and $\boldsymbol{\sigma} = \mathbf{1.0}$ ounces.

Mr. Kam, the owner, randomly selects a sample of 9 glasses and measures the amount of Guinness in each glass. Find the probability that the sample mean will be within 0.3 ounces of $\boldsymbol{\mu}$.

P[-0.9 < Z < 0.9] = 1 – 2(0.1841) = **0.6318**

- What if n = 16?   n = 49?   n = 100?

notes

Normal Example

Let $\quad X \sim N(5,4)$

$\quad\quad Y \sim N(4,1)$

Find P[Y > X]?

$P[Z > \frac{1}{\sqrt{5}}] = P[Z > 0.447] =$ **0.327**

Find P[Y > 2X]?

$P[Z > \frac{6}{\sqrt{17}}] = P[Z > 1.455] =$ **0.0728**

CLT Example

Suppose that achievement test scores of all college freshmen Illinois have a mean of 60, and variance 64. A random sample of 100 students from Pirdew had a mean score of 58. What is the probability that the mean score would be 58 or less if Pirdew is on par with other schools (i.e. if Pirdew has mean = 60).

$\mathrm{P}[\mathrm{Z} \leq -2.5] = 0.0062$ Yes.

notes