جلسه‌ی هفدهم آمار و احتمال مهندسی «دکتر شریفی زارچی - دانشگاه شریف»

فرض کنید میانگین (μ) طول عمر ابزار 60 سال باشد و SD را 10 سال فرض کنیم. احتمال 300 سال عمر کردن چقدر است؟

یک توزیع عجیب غریب داریم (ناشناخته) و فقط μ آن را داریم ($E(x)$). $P(X \geq a)$ چقدر است؟ [طبیعتاً X همواره از 0 بزرگتر است]

$$E(x) = \int_0^{+\infty} x f_X(x)\,dx = \int_0^{a} x f_X(x)\,dx + \int_a^{+\infty} x f_X(x)\,dx \geq \int_a^{+\infty} x f_X(x)\,dx \geq a \underbrace{\int_a^{+\infty} f_X(x)\,dx}_{\leq 1}$$

$$E(x) \geq a P(X \geq a) \Rightarrow \boxed{P(X \geq a) \leq \frac{E(x)}{a}}$$

نابرابری مارکوف

در سوال بالا $P(X \geq 300) \leq \frac{60}{300}$ یعنی نهایتاً 20 درصد احتمال داره بیشتر از 300 عمر کنه و از این دیگه بیشتر نمیشه.

هرچند دقت خوبی نیست اما یک قدم رو به جلو برداشتیم.

در قدم بعدی؛ واریانس رو هم داریم. حالا میشه نابرابری مارکوف رو با دقت بیشتری بدست آورد؟

$X \geq 0, \quad a > 0$

$$P(Y \geq b) \leq \frac{E(Y)}{b} \qquad P\left((X - E(x))^2 \geq a^2\right) \leq \frac{E\left((X - E(x))^2\right)}{a^2}$$

$Y = (X - E(x))^2$

$b = a^2$

$$\boxed{P(|X - E(x)| < a) \leq \frac{Var(x)}{a^2}}$$

نابرابری چبیشف

$$P(\underbrace{|X - 60|}_{>300} > \underbrace{a}_{240}) \leq \frac{\overset{\sim 10^2}{Var(x)}}{240^2} \Rightarrow P(|X - 60| > 240) \leq \frac{1}{576}$$

که خب خیلی بهتر شد :)

نامساوی چبیشف در حالتی که $X \geq 0$ رو نخواهیم به هیچ دردمون نمی‌خوره و باید دست‌کاریش کنیم.

$$\mathbb{P}(X \geq a) = \mathbb{P}(X+b \geq a+b) \leq \mathbb{P}\left((X+b)^2 \geq (a+b)^2\right)$$

$$\mathbb{P}(X \geq a) \leq \frac{E\left((X+b)^2\right)}{(a+b)^2} \longrightarrow \frac{\sigma^2+b^2}{(a+b)^2}$$

$$E\left((X+b)^2\right) = \underbrace{E(X^2)}_{\sigma^2} + b^2 + 2b\,\underbrace{E(X)}_{\text{فعلا فرض می‌کنیم که صفر هست.}}$$

if $b = \dfrac{\sigma^2}{a}$ $\Rightarrow$

$$\begin{cases} \mathbb{P}(X \geq a) \leq \dfrac{\sigma^2}{\sigma^2+a^2} \\ E(X) = 0 \\ a > 0 \end{cases}$$

برای اون 300 سال چیکار کنیم؟ باید مسئله‌مون رو به بالای ترسی بیاریم که نتیجه این نابرابری ازش استفاده کرد.

$$Y = X - \mu \longrightarrow \mathbb{P}(X-\mu \geq a) \leq \frac{\sigma^2}{a^2+\sigma^2} \Rightarrow \mathbb{P}(X-60 \geq 240) \leq \frac{1}{577}$$

یه ذره از اون دادامون بهتر شد و حتی دیگه قدر مطلق هم نداره و دقیقاً یک باند بالای به ما میده.

چرا $b = \dfrac{\sigma^2}{a}$؟ به ازای هر b ای اون رابطه‌ی بالا برقرار هست و حتی به ازای این مقدار به بهترین حالت خودش میرسه.

قانون اعداد بزرگ

فرض کنید می‌خواهیم میانگین قد تمام ایرانی‌ها را محاسبه کنیم. هر چقدر برای افراد بیشتر این میانگین رو محاسبه کنیم به تخمین بهتری می‌رسیم.
اما خب تا چه حد باید پیش بریم؟

$\varepsilon \geq 0$

$$\mathbb{P}\left(\frac{X_1+X_2+\cdots+X_n}{n} - \mu \geq \varepsilon\right) \longrightarrow 0$$

if $n \to \infty$

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_1+X_2+\cdots+X_n}{n}\right) = \mu$$ در واقع

فارسی طور بگیم: احتمال اینکه خطا از یک ε بزرگتر باشه در صورتی که $n \to \infty$ میل میکنه به سمت صفر. یعنی اگه آزمایش‌ها رو تکرار کنیم کم‌کم خطای ما کاهش پیدا میکنه. و اگر خیلی این کار رو تکرار کنیم خطا به ۰ میل میکنه.