Random Variables

Let X_i be random variables over $\mathbb{R}$.

Let $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ (empirical mean).

Let $\mu = \mathbb{E}[X]$

How far is $\bar{X}$ from μ?

$P(|\bar{X} - \mu| > t)$ should decay in t.

Come up with the <u>worst</u> probability distribution for decay rate wrt t.

Weak Law of Large Numbers

Let X_i be iid w/ $\mu = \mathbb{E}[X_i] < \infty$

$\sigma^2 = \mathbb{E}[(X_i-\mu)^2] < \infty$.

Then $\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mu$ in probability

pf: chebyshev's inequality

Central Limit Theorem

$$\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\sigma/\sqrt{n}} \longrightarrow N(0,1) \text{ in distribution}$$

Markov's Theorem

Let X be nonnegative r.v. w/ finite $\mathbb{E}(X)$.
For all $t > 0$

$$\mathbb{P}(X > t) \leq \frac{E[X]}{t}$$

Pf: $$\mathbb{E}(X) = \int_0^\infty x\, p(x)\,dx = \int_0^t x\, p(x)\,dx + \int_t^\infty x\, p(x)\,dx$$
$$\geq \int_t^\infty x\, p(x)\,dx \geq t \int_t^\infty p(x)\,dx = t\,\mathbb{P}(X > t)$$

Is Markov's Thm tight?
Q: is there a distribution st $P(X > t) = \frac{c}{t}$ $\forall\, t \geq T$?

No, such a dist would have density $\sim \frac{1}{t^2}$ which has ∞ avg.

Bound on Gaussian Tail

If $X \sim N(0,1)$, then $\mathbb{P}(|X|>\varepsilon) \leq \frac{2}{\varepsilon} e^{-\varepsilon^2/2}$

Proof: $$\mathbb{P}(|X|>\varepsilon) = 2\int_\varepsilon^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$\leq \frac{2}{\varepsilon} \int_\varepsilon^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} x\, dx$$

$$= \frac{2}{\sqrt{2\pi}\,\varepsilon} \int_{\varepsilon^2/2}^\infty e^{-u} du$$

$$= \sqrt{\frac{2}{\pi}} \frac{e^{-\varepsilon^2/2}}{\varepsilon} \leq \frac{e^{-\varepsilon^2/2}}{\varepsilon}$$

If $X_1 \cdots X_n \overset{iid}{\sim} N(0,1)$ then

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| > \varepsilon\right) \leq \frac{2}{\sqrt{n}\,\varepsilon} e^{-n\varepsilon^2/2}$$

Proof: $\sum_{i=1}^{n} X_i \sim N(0, n)$.

Let $Z = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(0, \frac{1}{n})$.

Note $\sqrt{n} Z \sim N(0,1)$

By Gaussian tail bound

$$\mathbb{P}(|Z| \geq \varepsilon) = \mathbb{P}(|\sqrt{n} Z| \geq \sqrt{n}\,\varepsilon) \leq \frac{2}{\sqrt{n}\,\varepsilon} e^{-n\varepsilon^2/2}$$

Chebyshev's Inequality

Let X have finite 1^{st} & 2^{nd} moments.

Let $\mu = \mathbb{E}(X)$. Let $\sigma^2 = Var(X)$.

$$\mathbb{P}(|X-\mu| > t) \leq \frac{\sigma^2}{t^2}, \quad P\left(\left|\frac{X-\mu}{\sigma}\right| > t\right) \leq \frac{1}{t^2}$$

pf: $$P(|X-\mu| > t) = P(|X-\mu|^2 > t^2) \leq \frac{\mathbb{E}(X-\mu)^2}{t^2} = \frac{\sigma^2}{t^2}.$$