Agenda:

1. Convergence of random variables
2. (Weak) law of large numbers
3. Central limit theorem

---

1. Convergence in probability

<u>Def</u> (convergence in probability to a constant): Let $\underbrace{X_1, X_2, \ldots}_{X_n}$ be an (infinite) sequence of random scalars and $c \in \mathbb{R}$ is a constant (c is not random), then $X_n \xrightarrow{p} c$ iff

$P(|X_n - c| > \epsilon) \to 0$ for any $\epsilon > 0$. (equiv) $\left(\lim_{n\to\infty} P(|X_n - c| > \epsilon) = 0\right)$

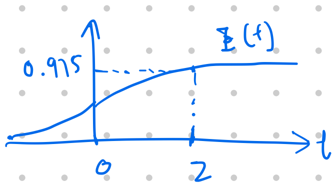
<u>Ex:</u> $X_n \sim N(0, \frac{1}{n})$

Let's check that $X_n \xrightarrow{p} 0$

$$P(|X_n - 0| > \epsilon) = P(X_n > \epsilon) + P(X_n < -\epsilon)$$
$$= \boxed{P(\underbrace{\sqrt{n}X_n}_{\sim N(0,1)} > \sqrt{n}\epsilon)} + \boxed{P(\sqrt{n}X_n < -\sqrt{n}\epsilon)}$$

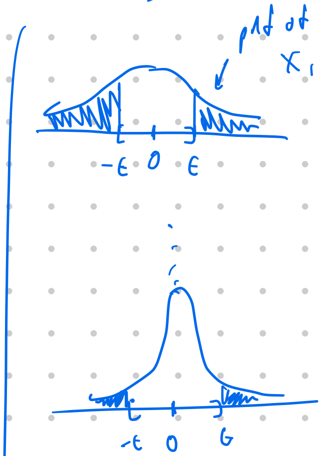
Define $\Phi(t)$ as the CDF of $N(0,1)$ random scalar

If $Z \sim N(0,1)$, then $\Phi(t) = P(Z \le t) = \int_{-\infty}^{t} \underbrace{\frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{x^2}{2}\right)}_{N(0,1)\text{ pdf}} dx$



$$= \underbrace{\boxed{1 - \Phi(\sqrt{n}\epsilon)}}_{\xrightarrow[n\to\infty]{} 1} + \underbrace{\boxed{\Phi(-\sqrt{n}\epsilon)}}_{\xrightarrow[n\to\infty]{} 0}$$

$\sqrt{n}\epsilon \xrightarrow[n\to\infty]{} \infty$, $-\sqrt{n}\epsilon \xrightarrow[n\to\infty]{} -\infty$



$\Rightarrow 1 - 1 + 0 = 0$

Def (Convergence in probability): $X_n \xrightarrow{p} X$ iff $X_n - X \xrightarrow{p} 0$

iff $P(\|X_n - X\|_2 > \epsilon) \to 0$

$\lim_{n\to\infty} P(\|X_n - X\|_2 > \epsilon) = 0$

for any fixed $\epsilon > 0$

<u>Ex:</u> law of large numbers (LLN)

Let $X_1, X_2, \ldots$ be an (independent & identically distributed) IID sequence of random scalars,

Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ as the sample mean of the first n random scalars

$X_1, \dots, X_n$

LLN claims: $\bar{X}_n \xrightarrow{P} \mu \triangleq E[X_1]$

Assume X_i has mean μ and variance σ^2

$E[\bar{X}_n] = \mu$, $Var[\bar{X}_n] = \sigma^2/n$

Markov's inequality: If X is a non-negative random scalar, then

$$P(X > t) \leq \frac{E[X]}{t} \quad \text{for any } t > 0$$

$$E[X] = E[X \mid X > t]\, P(X > t) + \underbrace{E[X \mid X \leq t]\, P(X \leq t)}_{\geq 0 \text{ b/c } X \text{ is non-negative}}$$

$$\geq E[X \mid X > t]\, P(X > t)$$

$$\geq t\, P(X > t)$$

---

pf of LLN:

$$P(|\bar{X}_n - \mu| \geq \epsilon) = P(|\bar{X}_n - \mu|^2 \geq \epsilon^2)$$

(Markov's inequality)
$$\leq \frac{E[|\bar{X}_n - \mu|^2]}{\epsilon^2}$$

$$= \frac{Var(\bar{X}_n)}{\epsilon^2}$$

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

$$\xrightarrow[n \to \infty]{} 0$$

$$E[\bar{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^n X_i\right] \quad \left(\text{def of } \bar{X}_n\right)$$

$$= \frac{1}{n}\sum_{i=1}^n E[X_i] \quad (\text{linearity of exp})$$

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

$$= \mu$$

$$Var[\bar{X}_n] = E[(\bar{X}_n - \mu)(\bar{X}_n - \mu)]$$

$$= E\left[\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right)^2\right]$$

$$= E\left[\left(\frac{1}{n}\sum_{i=1}^n (X_i - \mu)\right)^2\right]$$

$$= E\left[\frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n (X_i - \mu)(X_j - \mu)\right]$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^n E[(X_i - \mu)^2] + \sum_{i=1}^n \sum_{j \neq i} \cancel{E[(X_i - \mu) \times (X_j - \mu)]}^{\,0}\right)$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^n \sigma^2\right)$$

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

$$E[X] \overset{?}{=} E[X \mid X > t]\, P(X > t) + E[X \mid X \leq t]\, P(X \leq t)$$

$$\int x\, p(x)\, dx$$

$$= \int x\left(p(x, x > t) + p(x, x \leq t)\right) dx$$

$$= \int x\left(p(x \mid x > t)\, p(x > t)\right) dx + \left(\frac{p(x \mid x \leq t)\, x}{p(x \leq t)}\right) dx$$

$$= \underbrace{\boxed{\int x\, p(x \mid x > t)\, dx}}_{E[x \mid X > t]} \cdot p(x > t)$$

$$+ \underbrace{\boxed{\int x\, p(x \mid x \leq t)\, dx}}_{E[x \mid X \leq t]}\, p(x \leq t)$$

Convergence in distribution:

$\xrightarrow{d}$ ... $F_n(t) \to F(t)$ for any t such that

Def: $X_n \xrightarrow{d} X$ iff $F_n(t) \to F(t)$ for any t — F is continuous at t.

F_n is CDF of X_n: $F_n(t) = P(X_n \le t)$

F is CDF of X: $F(t) = P(X \le t)$

Alt def: $X_n \xrightarrow{d} X$ iff $E[g(X_n)] \to E[g(X)]$ for any bounded & cont. function g

---

Fact: If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{d} X$. (but reverse is generally not true)

Why? $F_n(t) = P(X_n \le t)$

$$= \underbrace{P(X_n \le t, X \le t+\epsilon)}_{\le P(X \le t+\epsilon)} + \underbrace{P(X_n \le t, X \ge t+\epsilon)}_{\le P(|X_n - X| \ge \epsilon)}$$

$$= F(t+\epsilon) + P(|X_n - X| \ge \epsilon)$$

Convergence in distribution to a constant:

If X_c is $\underbrace{\text{a constant}}_{P(X_c = c) = 1}$, then it's CDF is $F(t) = \begin{cases} 0 & \text{if } t < c \\ 1 & \text{if } t \ge c \end{cases}$

Claim: If $X_n \xrightarrow{d} c$, then $X_n \xrightarrow{P} c$.

$$P(|X_n - c| \ge \epsilon) = P(X_n < c-\epsilon) + P(X_n > c+\epsilon)$$

$$= \underbrace{F_n(c-\epsilon)}_{\to F(c-\epsilon) = 0} + 1 - \underbrace{F_n(c+\epsilon)}_{\to F(c+\epsilon) = 1}$$

$$= 0 + 1 - 1 = 0$$

Ex: central limit theorem:

Let $X_1, X_2, \ldots$ be IID random scalars, w/ mean μ & var. σ^2,

then $Z_n \triangleq \dfrac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z$, $Z \sim N(0,1)$