Last time: Moment generating function

Today: Law of large numbers
Convergence in probability

Law of large numbers

Motivation: Let $X$ be the outcome of a roll of a fair die. We have seen that its mean is

$$\mathbb{E}X = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \ldots + \frac{1}{6} \cdot 6 = \frac{7}{2} = 3.5.$$
What is the meaning of this number?

Answer: If we run the experiment $n$ times (roll $n$ dice) independently, then the average of the outcomes:

$$X_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

also called the sample mean

the 2nd die roll

should be close to 3.5.

Why is this?
Let's compute the mean and variance of $\bar{X}_n$:

\[
\rightarrow \quad E \bar{X}_n = E \left[ \frac{X_1 + \cdots + X_n}{n} \right] \\
\qquad = \frac{1}{n} E [X_1 + \cdots + X_n] \\
\qquad = \frac{1}{n} (E X_1 + \cdots + E X_n) = E X = 3.5.
\]

\[
\rightarrow \quad \text{Var} \bar{X}_n = \text{Var} \left( \frac{X_1 + \cdots + X_n}{n} \right) \\
\qquad = \frac{1}{n^2} \text{Var} (X_1 + \cdots + X_n) \\
\text{indep.} \quad \rightarrow \quad = \frac{1}{n^2} (\text{Var}(X_1) + \cdots + \text{Var}(X_n)) = \frac{1}{n} \cdot \text{Var} X = \frac{1}{n} \cdot \frac{35}{12}.
\]
When $n$ is large,

$\bar{X}_n$ has mean 3.5,

and small variance.

Thus, $\bar{X}_n$ is concentrated around 3.5.

That is, it's very likely that $\bar{X}_n$ is close to 3.5.
\[ n = 20 \]
$n = 100$
Theorem: (Weak law of large numbers – WLLN)

Let $X_1, X_2, \ldots$ be a sequence of independent r.v. with the same distribution (i.i.d. = independent and identically distributed), having mean $\mu$ and variance $\sigma^2$.

Define the sample mean:

$$\bar{X}_n = \frac{X_1 + \ldots + X_n}{n}.$$

Then

$$P\left(\left|\bar{X}_n - \mu\right| \geq \varepsilon\right) \xrightarrow{n \to \infty} 0 \quad \text{for any } \varepsilon > 0.$$
Proof: As before:

1. \( \mathbb{E} \bar{X}_n = \mu \)
2. \( \text{Var} \bar{X}_n = \frac{1}{n} \sigma^2 \)

Therefore, by Chebyshov's inequality,

\[
P( |\bar{X}_n - \mu| \geq \varepsilon ) \leq \frac{\text{Var} \bar{X}_n}{\varepsilon^2} = \frac{\sigma^2}{n \varepsilon^2} \xrightarrow{n \to \infty} 0
\]

since \( \frac{1}{n} \xrightarrow{n \to \infty} 0 \).
Convergence in probability

Let $X_1, X_2, \ldots$ be a sequence of r.v. defined on a common sample space, and let $X$ be another r.v. on the same sample space.

Definition: We say that $X_n$ converges to $X$ in probability if

$$\mathbb{P}( |X - X_n| \leq \varepsilon ) \xrightarrow{n \to \infty} 0$$

for any $\varepsilon > 0$. 
Example: If $X_1, X_2, \ldots$ are iid and $\bar{X}_n$ is the sample mean, then the WLLN states that $\bar{X}_n$ converges to $\mu$ in probability.

Notation: We write $X_n \xrightarrow{n \to \infty} X$.

Remark: $X_n \xrightarrow{n \to \infty} X$ means that $X_n = X + \eta_n$ where $\eta_n$ can be thought of as "noise".

$\mathbb{P}(|X_n| > \varepsilon) \xrightarrow{n \to \infty} 0$ for any $\varepsilon > 0$. 
Illustration: Let $X \sim \mathcal{N}(0,1)$

$Y_n \sim \mathcal{N}(0, \frac{1}{n^2})$ indep. of $X$.

Set $X_n = X + Y_n$.

Let's show that $X_n \xrightarrow{\text{prob}} X$.

Joint pdf of $X_n$ and $X$ is

$$f(x_n, x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{n}{\sqrt{2\pi}} e^{-\frac{n^2}{2} (x-x_n)^2}.$$
Observation: As \( n \) increases, the joint pdf concentrates around the diagonal \( x = x_n \).

This concentration is what convergence in probability means.

\[
\mathbb{P}( |X_n - X| \leq \varepsilon ) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy \\
\{ |x_n - x| \leq \varepsilon \}
\]

\[
= \int_{-\infty}^{\infty} \frac{n}{\sqrt{2\pi}} e^{-\frac{n^2}{2}y^2} \, dy = \mathbb{P}( |X_n| \leq \varepsilon )
\{ |y| \leq \varepsilon \}
\]

By Chebyshev's theorem, \( \lim_{n \to \infty} \mathbb{P}( |X_n| \leq \varepsilon ) = 1 \).