Law of Large Numbers.

\( X_1, ..., X_n \) - i.i.d. (independent, identically distributed)

\[
\overline{X} = \frac{X_1 + \ldots + X_n}{n} \rightarrow \text{as } n \rightarrow \infty, \overline{X}_1
\]

Can be used for functions of random variables as well:
Consider \( Y_i = r(X_1) \) - i.i.d.

\[
\overline{Y} = \frac{r(X_1) + \ldots + r(X_n)}{n} \rightarrow \text{as } n \rightarrow \infty, \overline{Y}_1 = \overline{E}\rightarrow Y_1
\]

Relevance for Statistics: Data points \( x_i \), as \( n \rightarrow \infty \).

The average converges to the unknown expected value of the distribution which often contains a lot (or all) of information about the distribution.

Example: Conduct a poll for 2 candidates:
\( p \in [0, 1] \) is what we're looking for
Poll: choose \( n \) people randomly: \( X_1, ..., X_n \)

\[
\overline{E}_X_1 = 1(p) + 0(1-p) = p \leftarrow \frac{X_1 + \ldots + X_n}{n} \text{ as } n \rightarrow \infty
\]

Other characteristics of distribution:

Moments of the distribution: for each integer, \( k \geq 1 \), kth moment \( \overline{E}X^k \)

Kth moment is defined only if \( \overline{E}|X|^k \) is finite

Moment generating function: consider a parameter \( y \in \mathbb{R} \). and define \( \phi(t) = \overline{E}e^{tx} \) where \( X \) is a random variable.

\( \phi(t) \) - m.g.f. of \( X \)

Taylor series of \( \phi(t) = \sum_{k=0}^{\infty} \frac{\phi^k(0)}{k!} t^k \)

Taylor series of \( \overline{E}e^{tx} = \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} \overline{E}X^k \)

\( \overline{E}X^k = \phi^k(0) \)

Example: Exponential distribution \( E(\alpha) \) with p.d.f. \( f(x) = \{ \alpha e^{-\alpha x}, x \geq 0; 0, x < 0 \} \)

Compute the moments:

\( \overline{E}X^k = \int_0^\infty x^k \alpha e^{-\alpha x} dx \) is a difficult integral.

Use the m.g.f.:

\[
\phi(t) = \overline{E}e^{tx} = \int_0^\infty e^{tx} \alpha e^{-\alpha x} dx = \int_0^\infty \alpha e^{(t-\alpha)x} dx
\]

(defined if \( t < \infty \) to keep the integral finite)
\[
\frac{\alpha e^{(t-\alpha)x}}{t-\alpha} \bigg|_{t=0}^{\infty} = 1 - \frac{\alpha}{t-\alpha} = \frac{\alpha}{\alpha - t} = \frac{1}{1-t/\alpha} = \sum_{k=0}^{\infty} \left(\frac{t}{\alpha}\right)^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} E^k
\]

Recall the formula for geometric series:
\[
\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad \text{when } k < 1
\]
\[
\frac{1}{\alpha^k} = \frac{E x^k}{k!} \rightarrow E x^k = \frac{k!}{\alpha^k}
\]

The moment generating function completely describes the distribution.
\[
E x^k = \int x^k f(x) dx
\]
If \(f(x)\) unknown, get a system of equations for \(f \rightarrow\) unique distribution for a set of moments.
M.g.f. uniquely determines the distribution.

\(X_1, X_2\) from \(E(\alpha), Y = X_1 + X_2\).
To find distribution of sum, we could use the convolution formula, but, it is easier to find the m.g.f. of sum \(Y\):
\[
E e^{tY} = E e^{t(X_1+X_2)} = E e^{tX_1} e^{tX_2} = E e^{tX_1} E e^{tX_2}
\]

Moment generating function of each:
\[
\frac{\alpha}{\alpha - t}
\]
For the sum:
\[
\frac{(\alpha/\alpha - t)^2}{\alpha - t}
\]
Consider the exponential distribution:
\[
E(\alpha) \sim X_1, E X = \frac{1}{\alpha}, f(x) = \{\alpha e^{-\alpha x}, x \geq 0; 0, x < 0\}
\]
This distribution describes the life span of quality products.
\(\alpha = 1/E, \) if \(\alpha\) small, life span is large.

**Median:**
\(m \in \mathbb{R}\) such that:
\[
\mathbb{P}(X \geq m) \geq \frac{1}{2}, \mathbb{P}(X \leq m) \geq \frac{1}{2}
\]
\[
\mathbb{P} = \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}
\]
(There are times in discrete distributions when the probability cannot ever equal exactly 0.5)
When you exclude the point itself: \(\mathbb{P}(X > m) \leq \frac{1}{2}\)
\(\mathbb{P}(X \leq m) + \mathbb{P}(X > m) = 1\)
The median is not always uniquely defined. Can be an interval where no point masses occur.
For a continuous distribution, you can define $P > m$ as equal to $\frac{1}{2}$.
But, there are still cases in which the median is not unique!

For a continuous distribution:

$$P(X \leq m) = P(X \geq m) = \frac{1}{2}$$

The average measures center of gravity, and is skewed easily by outliers.

The average will be pulled towards the tail of a p.d.f. relative to the median.

Mean: find $a \in \mathbb{R}$ such that $\mathbb{E}(X - a)^2$ is minimized over $a$.

$$\frac{\partial}{\partial a} \mathbb{E}(X - a)^2 = -2\mathbb{E}(X - a) = 0, \mathbb{E}X - a = 0 \rightarrow a = \mathbb{E}X$$

expectation - squared deviation is minimized.

Median: find $a \in \mathbb{R}$ such that $\mathbb{E}|X - a|$ is minimized.

$\mathbb{E}|X - a| \geq \mathbb{E}|X - m|$, where $m$ - median

$\mathbb{E}(|X - a| - |X - m|) \geq 0$

$f((x - a) - |x - m|)f(x)dx$
Need to look at each part:
1) \(a - x - (m - x) = a - m, x \leq m\)
2) \(x - a - (x - m) = m - a, x \geq m\)
3) \(a - x - (x + m) = a + m - 2x, m \leq x \leq a\)

The integral can now be simplified:

\[
\int (|x - a| - |x - m|) f(x) dx \geq \int_{-\infty}^{m} (a - m) f(x) dx + \int_{m}^{\infty} (m - a) f(x) dx = \\
= (a - m)(\int_{-\infty}^{m} f(x) dx - \int_{m}^{\infty} f(x) dx) = (a - m)(\mathbb{P}(X \leq m) - \mathbb{P}(X > m)) \geq 0
\]

As both \((a - m)\) and the difference in probabilities are positive.
The absolute deviation is minimized by the median.

** End of Lecture 18**