Recap:
1. Properties of a random sample
2. We have seen the definition of $t$-distribution, $F$-distribution and some of their properties

Order Statistic

There is an electronic device which needs 20 batteries. Each battery has the same distribution. Let the electronic device dies when $15$ batteries die. What is the distribution of the lifetime of the electronic device?

Let $X_1, \ldots, X_{20}$ be the lifetimes of 20 batteries. We know $X_1, \ldots, X_{20}$ is a random sample.

$$X_{(1)} = \min_{1 \leq i \leq 20} X_i, \quad X_{(2)} = \min \left\{ X_1, \ldots, X_{20} \setminus \{ X_{(1)} \} \right\}, \ldots$$

$$X_{(20)} = \max_{1 \leq i \leq 20} X_i$$

$X_{(1)} < X_{(2)} < \ldots < X_{(20)}$ These are known as the order statistics of the random sample.

$X_{(15)}$ is x.v. representing the time when 15 batteries die.
Marginal density of \(X_{(i)}\):

\[
 f_{X_{(i)}}(x) = \frac{n!}{(i-1)! (n-i)!} f_X(x) \left[ F_X(x) \right]^{i-1} \left[ 1 - F_X(x) \right]^{n-i}
\]

\(-\infty < x < \infty\)

\((^n_i) \rightarrow \) \(i\)th order statistic can be drawn

\(n \rightarrow (n-1)\)

\((^n_i) \rightarrow (i-1)\) order statistic below the
\n\(i\)th order statistic can be done in \((^n_{i-1})\) ways.

\(i\)th order stat. can take the value \(x\)
with "prob." (density) \(f_X(x)\).

And the \((i-1)\) order statistics below \(\Theta\) the \(i\)th one will have to be necessarily smaller than \(x\).

Each of them is \(< x\) with prob. \(F_X(x)\)

Then \((n-i)\) order statistic above the \(i\)th one will have to be \(> x\). Each of them is

\(> x\) with prob. \(1-F_X(x)\)

\[
 f_{X_{(i)}}(x) = \binom{n}{i}^{-1} \prod_{i=1}^{n-i} c_{i-1} f_X(x) \left[ F_X(x) \right]^{i-1} \left[ 1 - F_X(x) \right]^{n-i}
\]

\[= \frac{n!}{(i-1)! (n-i)!} f_X(x) \left[ F_X(x) \right]^{i-1} \left[ 1 - F_X(x) \right]^{n-i}\]
Joint density of $X(i, j)$

\[
\begin{align*}
&f_{X(i, j)}(x_1, x_2) = \frac{n!}{(i-1)! (j-i-1)! (n-j)!} \\
&\quad \times \left[ F_X(x_2) - F_X(x_1) \right]^{j-i-1} f_X(x_1)f_X(x_2), \\
&\quad x_1 \leq x_2
\end{align*}
\]

The $i$th order statistic can be chosen in $\binom{n}{i}$ ways. Once it has been chosen, the $j$th order statistic can be chosen in $\binom{n-1}{i}$ ways.

Once they are both chosen, the $(i-1)$ order statistics below the $i$th one can be chosen in $\binom{n-2}{i-1}$ ways. Once the $i$th, $j$th and all order statistics below the $i$th one are chosen, the $(n-j)$ order statistics above the $j$th one can be chosen in $\binom{n-2-(i-1)}{n-j}$ ways.

Now, the density of the $i$th order statistic being equal to $x_1$ is $f_X(x_1)$ and the density of the $j$th order statistic being $x_2$ is $f_X(x_2)$.

The prob. that $(i-1)$ order statistics below the $i$th one $\leq x_1$ is $\left[ F_X(x_1) \right]^{i-1}$.
the prob. that \((n-j)\) order statistics are above the \(j\)th one \(\geq x_2\) in \(\left[1 - F_X(x_2)\right]^{n-j}\)

Also, \((\binom{n}{j-i-1})\) order statistics which are
in between the \(i\)th and \(j\)th order statistic lie between \(x_1\) and \(x_2\) with prob.

\[
\left[ F_X(x_2) - F_X(x_1) \right]^{j-i-1}
\]

\[
\frac{\binom{n}{j-i-1}}{\binom{n}{i} \binom{n-i}{j-i-1}} F_X(x_1)^{i-1} \left[1 - F_X(x_2)\right]^{n-j}
\]

\[
F_X(x_2) - F_X(x_1)
\]

\(\text{Example:}\) \(x_1, \ldots, x_20\) are lifetimes of 20 batteries.

Let the lifetimes are all independently and identically distributed with density \(\text{Exp}(\lambda)\).

\(x_1, \ldots, x_{20} \sim \text{iid Exp}(\lambda)\).

\(\text{Goal!}\) \(\phi\) To find the distribution of \(\sum_{i=1}^{15} X_i\).

\[
F_{X(15)}(x) = \frac{20!}{14!5!} \left[\lambda \exp(-\lambda x)\right]^{14} \left[1 - \exp(-\lambda x)\right]^{5} \left[\exp(-\lambda x)\right]^{5}
\]

\(x > 0\)

\(\text{Example:}\) On a policy of five members
Example: Five members of a family are in an insurance policy. The policy says that they will receive huge amount of money of two people die. If the lifespan distributions of all people in the family are same. What is the distribution of the time when they receive the money?

Convergence of sequence of random variables

Let's say we have a random sample \( X_1, \ldots, X_n \) where \( n \) is very large. For example \( n = 20000 \). We are interested to know the distribution of some function of \( X_1, \ldots, X_{20000} \), let's say the distribution of \( \bar{X} = \frac{1}{20000} \sum_{i=1}^{20000} X_i \).

We might encounter a number of situations where the distribution of \( \bar{X} \) is difficult to treat analytically, but the distribution can be approximated by a well known distribution for large \( n \).
Convergence concepts

Convergence in probability

A sequence of random variables \( x_1, \ldots \) converges to a random variable \( x \) if

\[
\lim_{n \to \infty} P(|x_n - x| \geq \varepsilon) = 0
\]

\[
\frac{\varepsilon}{n!} \to 0
\]

Example: \( x_n \overset{D}{\to} \text{N}(0, \frac{1}{n}) \).

Qn: Does \( x_n \) converge in prob.? If yes, then to which \( \mu, \nu \)?

Guess: \( x = 0 \) w.p. 1.

\[
P(|x_n - x| \geq \varepsilon) = P(|x_n| \geq \varepsilon)
\]

\[
\leq \frac{E x_n}{\varepsilon^n} \quad \text{By Chebyshev's inequality}
\]

\[
= \frac{1}{n \varepsilon^n} \to 0 \quad \text{as} \quad n \to \infty.
\]

Qn: If \( x_1, \ldots \) converges in prob. to \( x \) (formally denoted by \( x_n \overset{p}{\to} x \)), does \( g(x_1), \ldots \) converge in prob. to \( g(x) \)?

If yes, under what condition on \( g \)?
\[ X_n \xrightarrow{p} x \Rightarrow g(X_n) \xrightarrow{p} g(x) \text{ when } g \text{ is a continuous function.} \]

Q:\[ X_n \xrightarrow{p} x \text{ and } Y_n \xrightarrow{p} y, \text{ where does } \]
\[ X_n + Y_n \xrightarrow{p} X + Y \text{ converge to some random variable in prob.?} \]

**Convergence in distribution**

\( X_1, X_2, \ldots \) is a sequence of random variables. This sequence is said to converge in distribution to a random variable \( X \) if for all continuity points \( x \) of \( F_X \),
\[ \lim_{n \to \infty} F_{X_n}(x) = F_X(x), \] where \( F_{X_n} \) and \( F_X \) are the cumulative density functions of \( X_n \) and \( X \).

**Example:** \( X_1, \ldots \) is a random sample from \( U(0,1) \). Does \( n(1-X(n)) \) converge in distribution?

We know \( X_n \sim U(0,1) \). Let \( Z_n = n(1-X(n)) \).

**Step 1:** Compute the C.D.F. of \( Z_n \).
\[ F_{Z_n}(z) = P(Z_n \leq z) = P(n(1-X(n)) \leq z) = P(X(n) \geq 1- \frac{z}{n}) \]
\[
\begin{align*}
1 - P(X_n < 1 - \frac{\alpha}{n}) &= 1 - P(X_1 < 1 - \frac{\alpha}{n}, X_2 < 1 - \frac{\alpha}{n}, \ldots, X_n < 1 - \frac{\alpha}{n}) \\
&= 1 - P(X_1 < 1 - \frac{\alpha}{n}) P(X_2 < 1 - \frac{\alpha}{n}) \ldots P(X_n < 1 - \frac{\alpha}{n}) \\
&= 1 - \left(1 - \frac{\alpha}{n}\right)^n \\
\text{as } x_1, \ldots, x_n \text{ are i.i.d.,} \\
\lim_{n \to \infty} F_{2n}(x) &= 1 - e^{-x} = F_Z(x) \text{ when } Z \sim \text{Exp}(1)
\end{align*}
\]

\( n(1-X_{(n)}) \) is a sequence of random variables that converges in distribution to \( \text{Exp}(1) \) random variable.

When \( X_n \) converges in distribution to \( X \) (formally written as \( X_n \xrightarrow{d} X \)), does it mean that \( X_n \xrightarrow{P} X \)?

This is not true.

**Counterexample:** \( X \) is a random variable s.t.
\[
P(X=0) = P(X=1) = \frac{1}{2}
\]

\( \Rightarrow X \) and \( 1-X \) have the same distribution.

Let's take \( X_n = X \ \forall n \).

\[\text{BOO} \]
clearly each $X_n$ has the same distribution as $1 - X$.

By default $X_n \overset{d}{\to} 1 - X$

\[ P(|X_n - (1-X)| > \frac{1}{2}) = P(|X - (1-X)| > \frac{1}{2}) \]

\[ = P(|2X-1| > \frac{1}{2}) \]

For both $X = 0$ and $1$, $|2X-1| = 1$

which means $P(|X_n - (1-X)| > \frac{1}{2}) = 1$

\[ \Rightarrow X_n \overset{p}{\to} 1 - X. \]

Hence convergence in distribution does not mean convergence in prob.

However: $X_n \overset{p}{\to} X \Rightarrow X_n \overset{d}{\to} X$.

Example: $X_n = \frac{1}{n}$ w.p. 1

$X = 0$ w.p. 1

$\frac{1}{n} \to 0$ We should expect $X_n \overset{d}{\to} X$ in this case.

However, $F_{X_n}(x) = \begin{cases} 0 & \text{if } x < \frac{1}{n} \\ 1 & \text{otherwise} \end{cases}$

$\lim_{n \to \infty} F_{X_n}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{otherwise} \end{cases}$

There is a discrepancy at $x = 0$ and if occurs as $F_X$ in not cont. at $x = 0$. 