Inequalities and Limits Part B

Outline

- Inequalities: Markov, Chebyshev, Jensen, Hölder
- Limits
  - of Probability: Weak Law of Large Numbers
  - of Events: Strong Law of Large Numbers
  - of Cdfs: Central Limit Theorem
Interest in Limits

Given a sequence \( \{X_1, X_2, \ldots \} \) of rvs, we are interested in the limit of the sequence.

Ex: Suppose we observe iid population \( \{X_1, X_2, \ldots \} \) with \( E(X_i) = \mu \) and \( V(X_i) = \sigma^2 \)

- Consider the mean of the sequence \( Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i \)
- For \( Y_n \), we have the moments: \( E(Y_n) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \mu \) and \( V(Y_n) = \frac{1}{n^2} \sum_{i=1}^{n} V(X_i) = \frac{\sigma^2}{n} \)
- Of interest is also probability of deviation of the mean of the sequence from the mean \( \mu \):
  - \( P(|Y_n - \mu| \leq \epsilon) \) for every \( n \) is tedious
  - Probability of deviation from the true mean, \( \lim_{n \to \infty} P(|Y_n - \mu| \leq \epsilon) \), is of interest

In general we are interested in \( \lim_{n \to \infty} P(X_n \in A) \) for a set \( A \). Some examples of \( A \).

- Ex: Let \( X_n \) be the number of (repeat) buyers in a market in week \( n \). \( X_1 = 200 \).

  \[
  X_{n+1} = \frac{1}{3} X_n + \begin{cases} 200 \ \text{with probability } \frac{2}{3} \\ 300 \ \text{with probability } \frac{1}{3} \\ 400 \ \text{with probability } \frac{1}{3} \end{cases}
  \]

  - Let \( a_1 = b_1 = 200 \), \( a_{n+1} = \frac{a_n}{3} + \frac{200}{3} \) and \( b_{n+1} = \frac{b_n}{3} + \frac{400}{3} \) so \( a_n = 100 \times \frac{3^{n-1}+1}{3^{n-1}} \) and \( b_n = 200 \)
  - \( P(X_n \in [a_n, b_n]) = 1 \) for every \( n \)
  - \( \lim_{n \to \infty} a_n = 100 \) and \( \lim_{n \to \infty} b_n = 200 \); \( \lim_{n \to \infty} P(X_n \in [100, 200]) = 1 \)

- Ex: For stock price \( X_n \), the probability that it doubles its current value: \( \lim_{n \to \infty} P(X_n \geq 2x_0) \).
Limit Examples

- **Ex:** For iid sequence \( \{X_1, X_2, \ldots \} \) with cdf \( F_X(x) = 1 - \exp(-\lambda x) \). Let \( Y_n = \min\{X_1, X_2, \ldots, X_n\} \), show for every \( \varepsilon > 0 \), \( \lim_{n \to \infty} P(|Y_n - 0| \leq \varepsilon) = 1 \).
  - For every \( n \), desired probability is \( P(Y_n \leq \varepsilon) = F_{Y_n}(\varepsilon) \).
  - \( F_{Y_n}(\varepsilon) = 1 - P(Y_n \geq \varepsilon) = 1 - P(X_1 \geq \varepsilon, \ldots, X_n \geq \varepsilon) = 1 - P(X_1 \geq \varepsilon)^n = 1 - \exp(-\lambda n\varepsilon) \).
  - \( \lim_{n \to \infty} F_{Y_n}(\varepsilon) = 1 \)
  - The minimum of the sequence gets closer to 0.

- **Ex:** Let iid \( \{X_1, X_2, \ldots\} \) be a sequence with mean \( \mu \) and vanishing variance \( \lim_{n \to \infty} V(X_n) = 0 \).
  Then for every \( \varepsilon > 0 \), the limit of the probability of staying in the neighborhood of the mean is
  \( \lim_{n \to \infty} P(|X_n - \mu| \leq \varepsilon) = 1 \)
  - To establish this use Chebyshev’s inequality \( P(|X_n - E(X_n)| \leq \varepsilon) \geq 1 - \frac{1}{\varepsilon^2} V(X_n) \) in the limit:
  \( \lim_{n \to \infty} P(|X_n - \mu| \leq \varepsilon) \geq \lim_{n \to \infty} \left( 1 - \frac{1}{\varepsilon^2} V(X_n) \right) = 1 \)

- **Can there be a generalization** of the last example so that the conclusion is \( \lim_{n \to \infty} P(|Y_n - \mu| \leq \varepsilon) = 1 \)
  for a special random variable \( Y_n \)?
Weak Law of Large Numbers: Let \( \{X_1, X_2, \ldots, X_n\} \) be an iid sequence with finite mean \( \mu \) and finite variance \( \sigma^2 \) and for the sample mean \( Y_n = \frac{\sum_{i=1}^{n} X_i}{n} \).

\[
\lim_{n \to \infty} P(|Y_n - \mu| \leq \epsilon) = 1.
\]

- The weak law: the sample mean \( Y_n \) remains in \((\mu - \epsilon, \mu + \epsilon)\) with probability converging to 1.
- In the weak law, the variance of \( Y_n \) is vanishing as it is \( \frac{\sigma^2}{n} \).
- Can we get the convergence without vanishing variance or without finite variance?

- Ex: \( X_n \) is a symmetric discrete rv with masses at only three points \([-n, 0, n]\), with the pmf

\[
P(X_n = x) = \begin{cases} 
\frac{1}{2n} & \text{if } |x| = n \\
1 - \frac{1}{n} & \text{if } x = 0 \\
0 & \text{otherwise}
\end{cases}
\]

- Each \( X_n \) is symmetric around 0, so \( E(X_n) = 0 \).
- The variance diverges: \( V(X_n) = E(X_n^2) = (-n)^2 \frac{1}{2n} + 0^2 \left(1 - \frac{1}{n}\right) + n^2 \frac{1}{2n} = n \)
- The weak law does NOT apply.
- But directly compute \( P(|X_n - 0| \leq \epsilon) : P(|X_n - 0| \leq \epsilon) = 1 - P(|X_n - 0| \geq \epsilon) = 1 - \frac{1}{n} \to 1 \).

Convergence is possible without vanishing variance. Is the weak law too weak for requiring too much?
Strong Law of Large Numbers: Let \( \{X_1, X_2, \ldots, X_n\} \) be an iid sequence with finite mean \( \mu \) and finite variance \( \sigma^2 \) and for the sample mean \( Y_n = \frac{\sum_{i=1}^{n} X_i}{n} \).

\[
P \left( \lim_{n \to \infty} |Y_n - \mu| \leq \epsilon \right) = 1.
\]

◆ The strong law is doubly stronger:
  1. The strong law has no condition of finite variance in its hypothesis: Less in hypothesis ⇒ Stronger
  2. The conclusion of the strong law implies the conclusion of the weak law. More in conclusion ⇒ Stronger

◆ 2. strength requires more discussion
  - The limits under consideration are
    \[
    \lim_{n \to \infty} P(|Y_n - \mu| \leq \epsilon) = \text{limit of numbers} \quad & \quad P \left( \lim_{n \to \infty} |Y_n - \mu| \leq \epsilon \right) = \text{limit of sets}
    \]
  - Are these different anyway?
    » The first limit above is over numbers; the second is over sets \( B_n = \{\omega \in \Omega: |Y_n(\omega) - \mu| \leq \epsilon\} \)
    » This answer is mathematically convincing but does not address the intuition
  - We need to understand \( \lim_{n \to \infty} P(B_n) \) versus \( P \left( \lim_{n \to \infty} B_n \right) \), where the difficulty is with the latter
Limits of Monotone Sets

- From earlier lectures

Increasing sequence

\[ A_4 = A_1 \cup A_2 \cup A_3 \cup A_4 \]

\[ P\left(\lim_{n \to \infty} A_n\right) = P\left(\bigcup_{i=1}^{\infty} A_i\right) \]

\[ P\left(\lim_{n \to \infty} A_n\right) = P\left(\bigcup_{i=1}^{\infty} A_i\right) \]

Decreasing sequence

\[ A_4 = A_1 \cap A_2 \cap A_3 \cap A_4 \]

\[ P\left(\lim_{n \to 4} A_n\right) = P\left(\bigcap_{i=1}^{4} A_i\right) \]

\[ P\left(\lim_{n \to 4} A_n\right) = P\left(\bigcap_{i=1}^{4} A_i\right) \]

- In the special case of increasing (decreasing) sets \( A_n \), there is a limiting set \( A \)
  
  \[ P\left(\lim_{n \to \infty} A_n\right) = P(A) \], because \( A \) is the limiting set

  \[ \lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) \] by the continuity of \( P \) for the limits of increasing (decreasing) sets

  Then \( \lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) \).
Limits of Nonmonotone Sets

- **lim sup** $A_n = \cap_{n=1}^{\infty} A_n^\cup$ where $A_n^\cup = \cup_{k=n}^{\infty} A_k$
  - Decreasing sequence: $A_n^\cup \downarrow$ as $n \uparrow$ because of fewer unions
  - $\limsup A_n = \text{limit of the decreasing sequence } A_n^\cup$
  - An element of $\limsup$ is present in an infinite number of sets

- **lim inf** $A_n = \cup_{n=1}^{\infty} A_n^\cap$ where $A_n^\cap = \cap_{k=n}^{\infty} A_k$
  - Increasing sequence: $A_n^\cap \uparrow$ as $n \uparrow$ because of fewer intersections
  - $\liminf A_n = \text{limit of the increasing sequence } A_n^\cap$
  - An element of $\liminf$ is absent from a finite number of sets

- Neither increasing sequence nor decreasing sequence of sets, instead of limit think of $\liminf$ and $\limsup$.

- You can get inspiration from $\liminf$ and $\limsup$ definition for the sequence of real numbers.

- **Infinite Presence** or **Finite Absence**, which is stricter?
Finite Absence is **Stricter Than** Infinite Presence

- Consider a particular \( \omega \in \Omega \),
  - \( \omega \in \lim \inf A_n \) if \( \omega \) is **absent from** finite number of \( A_n \)s
  - \( \omega \in \lim \sup A_n \) if \( \omega \) is **present in** infinite number of \( A_n \)s

- Ex: The difference between \( \lim \inf \) & \( \lim \sup \) is hard to explain to master and Ph.D. students. A professor decides to offer a lecture on this topic. He is patient and will offer this lecture every day and record the set of students attending in the set \( A_n \) for day \( n \geq 1 \). To avoid low attendance, he decides to
  - Limit the number of absences of a master student from the lectures by 10.
    » Absence from a finite number of lectures requires presence in infinite number of lectures
  - Limit the absence of Ph.D. students to at most 10% of the lectures
    » This is equivalent to requiring presence in at least 90% of the lectures
    » Presence in infinite number of lectures allows for absence in infinite number of lectures

\[
\lim \inf A_n = \{\text{Finite absences}\} = \{\text{Master stdnts}\} \subset \{\text{Master and Ph.D. stnts}\} = \lim \sup A_n
\]

Limited Absence Implies Infinite Presence

- Limitation to **finite absence** implies **infinite presence**
- **Infinite presence** does not imply finite absence
  - As infinite presence can occur simultaneously with infinite absence, e.g., a PhD student above
  - Infinite presence can be coupled with infinite absence
  - Infinite absence is more relaxed than **finite absence**
Example of Different Liminf and Limsup Sets

Consider the set
\[ A_n = I_{n \text{ odd}} \left( -1 - \frac{1}{n}, \frac{1}{n} \right) + I_{n \text{ even}} \left( -\frac{1}{n}, 1 + \frac{1}{n} \right) \] for \( n > 1 \)

For odd \( n \),
\[ A_n^\cap = \bigcap_{k=n}^\infty A_k = \left( -1 - \frac{1}{n}, \frac{1}{n} \right) \cap \left( -\frac{1}{n+1}, 1 + \frac{1}{n+1} \right) \cap \left( -1 - \frac{1}{n+2}, \frac{1}{n+2} \right) \cap \left( -\frac{1}{n+3}, 1 + \frac{1}{n+3} \right) \cap \ldots = \{0\} \]

- For odd \( n \), \( A_n^\cap = A_n \cap A_{n+1} = \left( -\frac{1}{n}, 1 + \frac{1}{n} \right) \cap \{0\} = \{0\} \)
- \( \lim \inf A_n = \bigcup_{n}^\infty A_n^\cap = \{0\} \).

For odd \( n \),
\[ A_n^\cup = \bigcup_{k=n}^\infty A_k = \left( -1 - \frac{1}{n}, \frac{1}{n} \right) \cup \left( -\frac{1}{n+1}, 1 + \frac{1}{n+1} \right) \cup \left( -1 - \frac{1}{n+2}, \frac{1}{n+2} \right) \cup \left( -\frac{1}{n+3}, 1 + \frac{1}{n+3} \right) \cup \ldots \]
\[ = \left( -1 - \frac{1}{n}, 1 + \frac{1}{n+1} \right) \cup \left( -1 - \frac{1}{n+2}, \frac{1}{n+3} \right) \cup \ldots \]

- For even \( n \), \( \bigcup_{k=n}^\infty A_k = A_n \cup A_{n+1} = \left( -\frac{1}{n}, 1 + \frac{1}{n} \right) \cup \left( -1 - \frac{1}{n+1}, 1 + \frac{1}{n+2} \right) = \left( -1 - \frac{1}{n+1}, 1 + \frac{1}{n} \right) \)
- \( [-1,1] \subseteq \lim \sup A_n \)
- \( 1 + \frac{1}{k} \notin \lim \sup A_n \) and \( 1 - \frac{1}{k} \notin \lim \sup A_n \) for any finite integer \( k \).
- \( \lim \sup A_n = [-1,1] \).

Hence, \( \lim \inf A_n = \{0\} \subseteq [-1,1] = \lim \sup A_n \)

- 1 is in \( A_n \) for even \( n \), so it is present in infinitely many sets and in \( \lim \sup \)
- 1 is absent from \( A_n \) for odd \( n \), so it is absent from infinitely many sets and from \( \lim \inf \)
- Replace 1 with any nonzero \( x \) with \(-1 \leq x \leq 1\).
From the Perspectives of Law of Large Numbers

- We have argued for \( \lim \inf A_n \subseteq \lim \sup A_n \). So \( P(\lim \inf A_n) \leq P(\lim \sup A_n) \)
- From limits of real numbers in analysis, we know: \( \lim \inf P(A_n) \leq \lim \sup P(A_n) \)
- From the definition of \( \lim \inf \) and \( \lim \sup \) over sets: \( P(\lim \inf A_n) \leq \lim \inf P(A_n) \leq \lim \sup P(A_n) \leq P(\lim \sup A_n) \)
- Putting together in a key inequality: \( P(\lim \inf A_n) \leq \lim \inf P(A_n) \leq \lim \sup P(A_n) \leq P(\lim \sup A_n) \)

- Ex: Let us go from the conclusion of the strong law to the weak law. Note \( Y_n \) is the sample mean.
  - \( B_n = \{ \omega \in \Omega: |Y_n - \mu| \leq \epsilon \} \)
  - The conclusion of the strong law: \( P(\lim_{n \to \infty} B_n) = 1 \), so \( \lim_{n \to \infty} B_n \) exists as a set or can even be taken as the sample space. These imply \( P(\lim \inf B_n) = P(\lim \sup B_n) = 1 \).
  - Inserting in the key inequality: \( 1 = P(\lim \inf B_n) \leq \lim \inf P(B_n) \leq \lim \sup P(B_n) \leq P(\lim \sup B_n) = 1 \)
  - Hence all inequalities become equality, in particular \( \lim \inf P(B_n) = \lim \sup P(B_n) = 1 \).
  - The sequence of real numbers \( \{P(B_n)\} \) has the limit of 1: \( 1 = \lim_{n \to \infty} P(B_n) = \lim_{n \to \infty} P(|Y_n - \mu| \leq \epsilon) \), which is the conclusion of the weak law.
  - Strong law is strong because
    - Conclusion of the strong law \( \Rightarrow \) Conclusion of the weak law.
    - Hypothesis of the weak law \( \Rightarrow \) Hypothesis of the strong law

- As hinted above, two convergence concepts:
  - \( P\left( \omega \in \Omega: \lim_{n \to \infty} |X_n(\omega) - X(\omega)| \leq \epsilon \right) = 1 \), we say \( X_n \) converges to \( X \) almost surely, we write \( X_n[\to as] X \)
  - \( \lim_{n \to \infty} P(\omega \in \Omega: |X_n(\omega) - X(\omega)| \leq \epsilon) = 1 \), we say \( X_n \) converges to \( X \) in probability, we write \( X_n[\to p] X \)

- Strong law has almost sure convergence and Weak law has convergence in probability

- Convergence almost surely (over sets) \( \Rightarrow \) Convergence in probability (over real numbers)
Observations on Convergence in Probability

Continuous functions preserve \([\rightarrow p]\): \(X_n[\rightarrow p]X_0\) and \(g\) continuous imply \(g(X_n)[\rightarrow p]g(X_0)\).

- To establish this, we start with the definition of continuity: For any \(\epsilon > 0\), there exists \(\delta > 0\) such that \(|x - x_0| < \delta\) implies \(|g(x) - g(x_0)| < \epsilon\). Thus,
  \[
  \{x: |x - x_0| < \delta\} \subseteq \{x: |g(x) - g(x_0)| < \epsilon\}
  \]
- This for random variables \(X_n\) and \(X_0\) implies
  \[
  P(|X_n - X_0| < \delta) \leq P(|g(X_n) - g(X_0)| < \epsilon).
  \]
- \(X_n[\rightarrow p]X_0\) requires \(\lim_{n \to \infty} P(|X_n - X_0| < \delta) = 1\). Combining this with the inequality above
  \[
  1 = \lim_{n \to \infty} P(|X_n - X_0| < \delta) \leq \lim_{n \to \infty} P(|g(X_n) - g(X_0)| < \epsilon) \quad \text{for any } \epsilon > 0.
  \]
- Hence, \(g(X_n)[\rightarrow p]g(X_0)\).
Observations on Convergence in Probability

Convergence of the Variance Estimator

Suppose that \( \{X_n\} \) is an iid sequence with \( \text{Normal}(0, \sigma^2) \). Consider \( Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \) and show that \( Y_n \to_p \sigma^2 \). Significance: \( Y_n \) can be used as the estimator of the variance \( \sigma^2 \).

Find the distribution of \( X_i^2 \). For \( u \geq 0 \),

\[
P(X_i^2 \leq u) = P(-\sqrt{u} \leq X_i \leq \sqrt{u}) = P(X_i \leq \sqrt{u}) - P(X_i \leq -\sqrt{u}) = P(X_i \leq \sqrt{u}) - P(X_i \geq \sqrt{u}) = P(X_i \leq \sqrt{u}) - \left(1 - P(X_i \leq \sqrt{u})\right) = 2P(X_i \leq \sqrt{u}) - 1,
\]

due to the symmetry of \( \text{Normal}(0, \sigma^2) \). The pdf of \( X_i^2 \) is

\[
f_{X_i^2}(u) = 2 \frac{1}{2\sqrt{u}} f_{\text{Normal}(0, \sigma^2)}(\sqrt{u}) = \frac{1}{\sqrt{u}2\pi\sigma^2} \exp\left(-\frac{u}{2\sigma^2}\right) \text{ for } u \geq 0.
\]

Recall that \( \Gamma(a = 1/2, \lambda = 1/(2\sigma^2)) \) has the pdf

\[
f_{\Gamma(a=1/2, \lambda=1/(2\sigma^2))}(x) = \left(\frac{1}{2\sigma^2}\right) \exp\left(-\frac{x}{2\sigma^2}\right) \left(\frac{1}{2\sigma^2}\right)^{1/2-1} \Gamma\left(\frac{1}{2}\right) = \frac{1}{\sqrt{\pi\sqrt{2}\sigma^2}\sqrt{x}} \exp\left(-\frac{x}{2\sigma^2}\right) \text{ for } x \geq 0.
\]

Hence, \( X_i^2 \sim \Gamma(a = 1/2, \lambda = 1/(2\sigma^2)) \) and \( \sum_{i=1}^{n} X_i^2 \sim \Gamma(a=n/2, \lambda=1/(2\sigma^2)) \).

\[
E(\sum_{i=1}^{n} X_i^2) = \frac{n/2}{1/(2\sigma^2)} = n\sigma^2 \text{ and } V(\sum_{i=1}^{n} X_i^2) = \frac{n/2}{1/(4\sigma^4)} = 2n\sigma^4 \text{ which imply}
\]

\[
E(Y_n) = E\left(\frac{1}{n} \sum_{i=1}^{n} X_i^2\right) = \sigma^2 \text{ and } V(Y_n) = V\left(\frac{1}{n} \sum_{i=1}^{n} X_i^2\right) = 2\sigma^4/n.
\]

\( Y_n \) has vanishing variance, so it converges to its mean in probability: \( Y_n \to_p \sigma^2 \).
Convergence in Distribution

- Suppose that each random variable $X_n$ in the sequence $\{X_n\}$ has a cumulative distribution function $F_n$ and $X$ has a cumulative distribution function $F$, if
  \[ \lim_{n \to \infty} F_{X_n}(a) = F_X(a) \text{ for each } a \text{ except for discontinuity points of } F_X, \]
  then we say $X_n$ converges to $X$ in distribution and write $X_n \xrightarrow{d} X$.

- Ex: Let $X_n \sim Beta(\alpha_n, \lambda_n)$ for $\alpha_n, \lambda_n \geq 0$ and $\alpha_n \to \alpha, \lambda_n \to \lambda$. Then $X_n \xrightarrow{d} X$ for $X \sim Beta(\alpha, \lambda)$.
  - To check this we can consider the pdf of $Beta(\alpha_n, \lambda_n)$ as $F_{X_n}(a) = \int_0^a f_{X_n}(x)dx$.
    \[ f_{X_n}(x) = \frac{x^{\alpha_n-1}(1-x)^{\lambda_n-1}\Gamma(\alpha_n+\lambda_n)}{\Gamma(\alpha_n)\Gamma(\lambda_n)} \]
  - This pdf converges to the pdf of $Beta(\alpha, \lambda)$ provided that $\Gamma(\alpha_n) \to \Gamma(\alpha), \Gamma(\lambda_n) \to \Gamma(\lambda)$ and $\Gamma(\alpha_n + \lambda_n) \to \Gamma(\alpha + \lambda)$. The last three convergences hold because $\Gamma(x)$ is a continuous function for $x \geq 0$.

- Ex: Suppose $\{X_1, ..., X_n\}$ is an iid sequence with $X_i \sim Expo(\lambda)$. Let $Y_n = \max\{X_1, X_2, ..., X_n\}$ and find the limiting distribution of the shifted version of $Y_n$: $Z_n = Y_n - \ln(n)/\lambda$
  - $Z_n \in [\ln(n)/\lambda, \infty)$. Since $Y_n$ is an order statistics, we have $P(Y_n \leq y) = (1 - e^{-\lambda y})^n$. Then
    \[ F_{Z_n}(z) = P(Z_n \leq z) = P\left(Y_n \leq z + \frac{\ln n}{\lambda}\right) = \left(1 - e^{-\lambda \left(z + \frac{\ln n}{\lambda}\right)}\right)^n = \left(1 - \frac{e^{-\lambda z}}{e^{\ln n}}\right)^n \]
  - Hence, $\lim_{n \to \infty} F_{Z_n}(z) = e^{-e^{-\lambda z}} = F_Z(z)$ for $-\infty < z < \infty$. $Z_n \xrightarrow{d} Z$, where
    » $Z$ has double exponential distribution or Gumbel distribution with parameter $\lambda$.
    » Gumbel & his contemporaries studied limiting distributions of extreme values.

- Gumbel investigated the water levels in rivers and flood patterns.
- Gumbel distribution is later used in consumer choice theory to model the maximum utility obtained from a set of choices and then to end up with the celebrated multinomial choice model of marketing.
Convergence in Distribution - Examples

- Ex: Suppose \( \{X_1, \ldots, X_n\} \) is an iid sequence with \( X_i \sim \text{Pareto}(\alpha, x_u) \) for \( x_u, \alpha \geq 0 \). Let \( Y_n = \max\{X_1, X_2, \ldots, X_n\} \) and find the limiting distribution of the scaled version of \( Y_n \): \( Z_n = \frac{Y_n}{n^{1/\alpha}} \).

  - \( Z_n \in \left[ x_u n^{-1/\alpha}, \infty \right) \). Since \( Y_n \) is an order statistics, we have \( P(Y_n \leq y) = \left(1 - \left(\frac{x_u}{y}\right)^\alpha\right)^n \). Then

    \[
    F_{Z_n}(z) = P(Z_n \leq z) = P(Y_n \leq n^{1/\alpha} z) = \left(1 - \left(\frac{x_u}{n^{1/\alpha} z}\right)^\alpha\right)^n = \left(1 - \frac{z}{x_u} n \right)^\alpha \]

    - Hence, \( \lim_{n \to \infty} F_{Z_n}(z) = e^{-(z/x_u)^{-\alpha}} = e^{-(x_u/z)^\alpha} =: F_z(z) \) for \( 0 < z < \infty \), which is a Weibull type distribution.

  - To better see why, consider \( W = \frac{1}{Z} \) as a new random variable. Then for \( w \geq 0 \)

    \[
    P(W \leq w) = P\left(\frac{1}{Z} \leq w\right) = P\left(Z \geq \frac{1}{w}\right) = 1 - P\left(Z \leq \frac{1}{w}\right) = 1 - e^{-\left(\frac{w}{x_u}\right)^\alpha} = 1 - \exp\left(-\left(\frac{w}{1/x_u}\right)^\alpha\right),
    \]

    where \( 1/x_u \) and \( \alpha \) are known respectively as scale and shape parameters of a Weibull distribution. When \( \alpha = 1 \), Weibull distribution becomes exponential distribution with mean \( 1/x_u \).

- Ex: Can continuous distributions \( \rightarrow d \) to a discrete distribution? Yes, consider the following:

\[
F_n(x) = \begin{cases} \frac{(2x)^n}{2} & \text{for } 0 \leq x < \frac{1}{2} \\ \frac{1}{2} + \frac{(2x-1)^n}{2} & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases}
\]

\[
F(x) = \begin{cases} 0 & \text{for } x < \frac{1}{2} \\ \frac{1}{2} & \text{for } \frac{1}{2} \leq x < 1 \\ 1 & \text{for } 1 \leq x \end{cases}
\]

\( F_n(0) = 0, F_n(1) = 1 \) and \( F_n \) is continuous everywhere including \( x = 1/2 \)

\( F \) is discontinuous; has masses of 1/2 at \( x = 1/2 \) and \( x = 1 \):
Central Limit Theorem: Limits of Cumulative Distributions

- **Central Limit Theorem:** Let \( \{X_1, X_2, \ldots, X_n\} \) be an iid sequence with finite mean \( \mu \) and variance \( \sigma^2 \), consider the shifted and scaled sample mean

\[
Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}
\]

Then \( Z_n \overset{d}{\to} \text{Normal}(0,1) \).

- Ex: A laser pointer uses a battery with uncertain lifetime. Although the distribution of the lifetime is not known, the mean and standard deviation are found respectively to be 10 hours and 2 hours. An instructor teaches 9 courses in a year and uses the pointer for 180 hours in total. What is the approximate probability that more than 19 batteries are needed in that year.

  - Let \( X_i \) be the lifetime of \( i \)th battery and let \( X^{19} = X_1 + X_2 + \cdots + X_{19} \) be time that 19 batteries last. The desired probability is \( P(X^{19} < 180) \). We can use the central limit theorem to approximate this as follows:

\[
P(X^{19} < 180) = P\left(\frac{X^{19} - n\mu}{\sqrt{n}\sigma} < \frac{180 - n\mu}{\sqrt{n}\sigma}\right) = P\left(\frac{X^{19} - 19(10)}{\sqrt{19} \cdot 2} < \frac{180 - 19(10)}{\sqrt{19} \cdot 2}\right)
\]

\[
\rightarrow P\left(\text{Normal}(0,1) \leq -\frac{5}{\sqrt{19}}\right) = P(\text{Normal}(0,1) \leq -1.147) = \text{Normdist}(-1.147,0,1,1) = 0.1257,
\]

where Normdist is a standard Excel function. So 19 batteries suffice for a year with a probability of 87%.
Summary

◆ Inequalities: Markov, Chebyshev, Jensen, Hölder
◆ Limits:
  – of Probability: Weak Law of Large Numbers
  – of Events: Strong Law of Large Numbers
  – of Cdfs: Central Limit Theorem

Let $\{X_i\}$ be an iid sequence with finite mean $\mu$ & variance $\sigma^2$ and consider $Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.

<table>
<thead>
<tr>
<th>Results</th>
<th>$Y_n \xrightarrow{p} \mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak law of large numbers</td>
<td>$Y_n \xrightarrow{as} \mu$</td>
</tr>
<tr>
<td>Strong law of large numbers</td>
<td>$Y_n \xrightarrow{d} \text{Normal}(\mu, \sigma^2/n)$</td>
</tr>
</tbody>
</table>