18.175: Lecture 14

Weak convergence and characteristic functions

Scott Sheffield

MIT

Outline

Weak convergence

Characteristic functions

Outline

Weak convergence

Characteristic functions

Convergence results

- **Theorem:** If $F_n \to F_\infty$, then we can find corresponding random variables Y_n on a common measure space so that $Y_n \to Y_\infty$ almost surely.

Convergence results

- **Theorem:** If $F_n \to F_\infty$, then we can find corresponding random variables Y_n on a common measure space so that $Y_n \to Y_\infty$ almost surely.
- **Proof idea:** Take $\Omega = (0, 1)$ and $Y_n = \sup\{y : F_n(y) < x\}$.

Convergence results

- **Theorem:** If $F_n \to F_\infty$, then we can find corresponding random variables Y_n on a common measure space so that $Y_n \to Y_\infty$ almost surely.
- **Proof idea:** Take $\Omega = (0, 1)$ and $Y_n = \sup\{y : F_n(y) < x\}$.
- **Theorem:** $X_n \implies X_\infty$ if and only if for every bounded continuous g we have $Eg(X_n) \to Eg(X_\infty)$.

Convergence results

- **Theorem:** If $F_n \to F_\infty$, then we can find corresponding random variables Y_n on a common measure space so that $Y_n \to Y_\infty$ almost surely.
- **Proof idea:** Take $\Omega = (0,1)$ and $Y_n = \sup\{y : F_n(y) < x\}$.
- **Theorem:** $X_n \implies X_\infty$ if and only if for every bounded continuous g we have $Eg(X_n) \to Eg(X_\infty)$.
- **Proof idea:** Define X_n on common sample space so converge a.s., use bounded convergence theorem.

Convergence results

- **Theorem:** If $F_n \to F_\infty$, then we can find corresponding random variables Y_n on a common measure space so that $Y_n \to Y_\infty$ almost surely.
- **Proof idea:** Take $\Omega = (0,1)$ and $Y_n = \sup\{y : F_n(y) < x\}$.
- **Theorem:** $X_n \implies X_\infty$ if and only if for every bounded continuous g we have $Eg(X_n) \to Eg(X_\infty)$.
- **Proof idea:** Define X_n on common sample space so converge a.s., use bounded convergence theorem.
- **Theorem:** Suppose g is measurable and its set of discontinuity points has μ_X measure zero. Then $X_n \implies X_\infty$ implies $g(X_n) \implies g(X)$.

Convergence results

- **Theorem:** If $F_n \to F_\infty$, then we can find corresponding random variables Y_n on a common measure space so that $Y_n \to Y_\infty$ almost surely.
- **Proof idea:** Take $\Omega = (0,1)$ and $Y_n = \sup\{y : F_n(y) < x\}$.
- **Theorem:** $X_n \implies X_\infty$ if and only if for every bounded continuous g we have $Eg(X_n) \to Eg(X_\infty)$.
- **Proof idea:** Define X_n on common sample space so converge a.s., use bounded convergence theorem.
- **Theorem:** Suppose g is measurable and its set of discontinuity points has μ_X measure zero. Then $X_n \implies X_\infty$ implies $g(X_n) \implies g(X)$.
- **Proof idea:** Define X_n on common sample space so converge a.s., use bounded convergence theorem.

Compactness

- **Theorem:** Every sequence F_n of distribution has subsequence converging to right continuous nondecreasing F so that $\lim F_{n(k)}(y) = F(y)$ at all continuity points of F.

Compactness

- **Theorem:** Every sequence F_n of distribution has subsequence converging to right continuous nondecreasing F so that $\lim F_{n(k)}(y) = F(y)$ at all continuity points of F.
- Limit may not be a distribution function.

Compactness

- **Theorem:** Every sequence F_n of distribution has subsequence converging to right continuous nondecreasing F so that $\lim F_{n(k)}(y) = F(y)$ at all continuity points of F.
- Limit may not be a distribution function.
- Need a "tightness" assumption to make that the case. Say μ_n are **tight** if for every ϵ we can find an M so that $\mu_n[-M, M] < \epsilon$ for all n. Define tightness analogously for corresponding real random variables or distributions functions.

Compactness

- **Theorem:** Every sequence F_n of distribution has subsequence converging to right continuous nondecreasing F so that $\lim F_{n(k)}(y) = F(y)$ at all continuity points of F.
- Limit may not be a distribution function.
- Need a "tightness" assumption to make that the case. Say μ_n are **tight** if for every ϵ we can find an M so that $\mu_n[-M, M] < \epsilon$ for all n. Define tightness analogously for corresponding real random variables or distributions functions.
- **Theorem:** Every subsequential limit of the F_n above is the distribution function of a probability measure if and only if the F_n are tight.

Total variation norm

- If we have two probability measures μ and ν we define the **total variation distance** between them is $||\mu - \nu|| := \sup_B |\mu(B) - \nu(B)|$.

Total variation norm

- If we have two probability measures μ and ν we define the **total variation distance** between them is $||\mu - \nu|| := \sup_B |\mu(B) - \nu(B)|$.
- Intuitively, it two measures are close in the total variation sense, then (most of the time) a sample from one measure looks like a sample from the other.

Total variation norm

- If we have two probability measures μ and ν we define the **total variation distance** between them is $||\mu - \nu|| := \sup_B |\mu(B) - \nu(B)|$.
- Intuitively, it two measures are close in the total variation sense, then (most of the time) a sample from one measure looks like a sample from the other.
- Corresponds to L_1 distance between density functions when these exist.

Total variation norm

- If we have two probability measures μ and ν we define the **total variation distance** between them is $||\mu - \nu|| := \sup_B |\mu(B) - \nu(B)|$.
- Intuitively, it two measures are close in the total variation sense, then (most of the time) a sample from one measure looks like a sample from the other.
- Corresponds to L_1 distance between density functions when these exist.
- Convergence in total variation norm is much stronger than weak convergence. Discrete uniform random variable U_n on $(1/n, 2/n, 3/n, \ldots, n/n)$ converges weakly to uniform random variable U on $[0, 1]$. But total variation distance between U_n and U is 1 for all n.

Outline

Weak convergence

Characteristic functions

Outline

Weak convergence

Characteristic functions

Characteristic functions

- Let X be a random variable.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- Recall that by definition $e^{it} = \cos(t) + i\sin(t)$.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- Recall that by definition $e^{it} = \cos(t) + i\sin(t)$.
- Characteristic function ϕ_X similar to moment generating function M_X.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- Recall that by definition $e^{it} = \cos(t) + i\sin(t)$.
- Characteristic function ϕ_X similar to moment generating function M_X.
- $\phi_{X+Y} = \phi_X \phi_Y$, just as $M_{X+Y} = M_X M_Y$, if X and Y are independent.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- Recall that by definition $e^{it} = \cos(t) + i\sin(t)$.
- Characteristic function ϕ_X similar to moment generating function M_X.
- $\phi_{X+Y} = \phi_X \phi_Y$, just as $M_{X+Y} = M_X M_Y$, if X and Y are independent.
- And $\phi_{aX}(t) = \phi_X(at)$ just as $M_{aX}(t) = M_X(at)$.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- Recall that by definition $e^{it} = \cos(t) + i\sin(t)$.
- Characteristic function ϕ_X similar to moment generating function M_X.
- $\phi_{X+Y} = \phi_X \phi_Y$, just as $M_{X+Y} = M_X M_Y$, if X and Y are independent.
- And $\phi_{aX}(t) = \phi_X(at)$ just as $M_{aX}(t) = M_X(at)$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.

Characteristic functions

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- Recall that by definition $e^{it} = \cos(t) + i\sin(t)$.
- Characteristic function ϕ_X similar to moment generating function M_X.
- $\phi_{X+Y} = \phi_X\phi_Y$, just as $M_{X+Y} = M_X M_Y$, if X and Y are independent.
- And $\phi_{aX}(t) = \phi_X(at)$ just as $M_{aX}(t) = M_X(at)$.
- And if X has an mth moment then $E[X^m] = i^m\phi_X^{(m)}(0)$.
- Characteristic functions are well defined at all t for all random variables X.

Characteristic function properties

- $\phi(0) = 1$

Characteristic function properties

- $\phi(0) = 1$
- $\phi(-t) = \overline{\phi(t)}$

Characteristic function properties

- $\phi(0) = 1$
- $\phi(-t) = \overline{\phi(t)}$
- $|\phi(t)| = |Ee^{itX}| \leq E|e^{itX}| = 1.$

Characteristic function properties

- $\phi(0) = 1$
- $\phi(-t) = \overline{\phi(t)}$
- $|\phi(t)| = |Ee^{itX}| \leq E|e^{itX}| = 1.$
- $|\phi(t+h) - \phi(t)| \leq E|e^{ihX} - 1|$, so $\phi(t)$ uniformly continuous on $(-\infty, \infty)$

Characteristic function properties

- $\phi(0) = 1$
- $\phi(-t) = \overline{\phi(t)}$
- $|\phi(t)| = |Ee^{itX}| \leq E|e^{itX}| = 1$.
- $|\phi(t+h) - \phi(t)| \leq E|e^{ihX} - 1|$, so $\phi(t)$ uniformly continuous on $(-\infty, \infty)$
- $Ee^{it(aX+b)} = e^{itb}\phi(at)$

Characteristic function examples

- **Coin:** If $P(X = 1) = P(X = -1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.

Characteristic function examples

- **Coin:** If $P(X=1) = P(X=-1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?

Characteristic function examples

- **Coin:** If $P(X=1) = P(X=-1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?
- **Poisson:** If X is Poisson with parameter λ then $\phi_X(t) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k e^{itk}}{k!} = \exp(\lambda(e^{it} - 1))$.

Characteristic function examples

- **Coin:** If $P(X=1) = P(X=-1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?
- **Poisson:** If X is Poisson with parameter λ then $\phi_X(t) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k e^{itk}}{k!} = \exp(\lambda(e^{it} - 1))$.
- Why does doubling λ amount to squaring ϕ_X?

Characteristic function examples

- **Coin:** If $P(X=1) = P(X=-1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?
- **Poisson:** If X is Poisson with parameter λ then $\phi_X(t) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k e^{itk}}{k!} = \exp(\lambda(e^{it} - 1))$.
- Why does doubling λ amount to squaring ϕ_X?
- **Normal:** If X is standard normal, then $\phi_X(t) = e^{-t^2/2}$.

Characteristic function examples

- **Coin:** If $P(X=1)=P(X=-1)=1/2$ then $\phi_X(t)=(e^{it}+e^{-it})/2=\cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?
- **Poisson:** If X is Poisson with parameter λ then $\phi_X(t)=\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k e^{itk}}{k!}=\exp(\lambda(e^{it}-1))$.
- Why does doubling λ amount to squaring ϕ_X?
- **Normal:** If X is standard normal, then $\phi_X(t)=e^{-t^2/2}$.
- Is ϕ_X always real when the law of X is symmetric about zero?

Characteristic function examples

- **Coin:** If $P(X = 1) = P(X = -1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?
- **Poisson:** If X is Poisson with parameter λ then $\phi_X(t) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k e^{itk}}{k!} = \exp(\lambda(e^{it} - 1))$.
- Why does doubling λ amount to squaring ϕ_X?
- **Normal:** If X is standard normal, then $\phi_X(t) = e^{-t^2/2}$.
- Is ϕ_X always real when the law of X is symmetric about zero?
- **Exponential:** If X is standard exponential (density e^{-x} on $(0, \infty)$) then $\phi_X(t) = 1/(1 - it)$.

Characteristic function examples

- **Coin:** If $P(X = 1) = P(X = -1) = 1/2$ then $\phi_X(t) = (e^{it} + e^{-it})/2 = \cos t$.
- That's periodic. Do we always have periodicity if X is a random integer?
- **Poisson:** If X is Poisson with parameter λ then $\phi_X(t) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k e^{itk}}{k!} = \exp(\lambda(e^{it} - 1))$.
- Why does doubling λ amount to squaring ϕ_X?
- **Normal:** If X is standard normal, then $\phi_X(t) = e^{-t^2/2}$.
- Is ϕ_X always real when the law of X is symmetric about zero?
- **Exponential:** If X is standard exponential (density e^{-x} on $(0, \infty)$) then $\phi_X(t) = 1/(1 - it)$.
- **Bilateral exponential:** if $f_X(t) = e^{-|x|}/2$ on $\mathbb{R}$ then $\phi_X(t) = 1/(1 + t^2)$. Use linearity of $f_X \to \phi_X$.