ECON 7710

Econometrics I

Lecture notes 1.

Elements of set theory:

- Set is arbitrary collection of items; subset of items A of set S is called its subset, denote $A \subset S$
- The set with no elements is called an empty set (denoted $\emptyset$)
- In probability usually deal with sets of "outcomes" (subsets of sample space Ω) called events
- Set operations $A \cup B = \{x : x \in A \text{ or } x \in B\}$ $A \cap B = \{x : x \in A \text{ and } x \in B\}$
  1. $A \cup B = B \cup A$, $A \cap B = B \cap A$
  2. $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$
  3. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \setminus B = \{x : x \in A, x \notin B\}$
- For each $A \subset \Omega$, complement $A^c = \Omega \setminus A$
- Sigma-algebra defines "order" of sets
  1. $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$.
  2. $A_1, A_2, \ldots, \in \mathcal{F} \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.
  3. $\emptyset \in \mathcal{F}$.

  Question: Show that the above conditions imply $\Omega \in \mathcal{F}$ and $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.
- Sigma algebra induced by all segments of real line is called Borel sigma-algebra
- $(\Omega, \mathcal{F})$ is called measurable space

- Function $f : A \mapsto B$ with $\mathcal{A}$- sigma algebra on A and $\mathcal{B}$- sigma algebra on B such that for any $S_a \in \mathcal{B}$, $f^{-1}(S_a) \in \mathcal{A}$ is called measurable function

Probability Space: $(\Omega, \mathcal{F}, P)$

- Random variable is a measurable function on algebra of events
- Probability measure P is a set function on $\mathcal{F}$ such that
  1. $\forall A \in \mathcal{F}$, $P(A) \geq 0$
  2. $P(\Omega) = 1$
  3. $\forall \{A_i\}_{i=1}^{\infty}$ such that $A_i \cap A_j = \emptyset$, $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$
- Properties of probability measure
  1. $P(\emptyset) = 0$, $P(A) \leq 1$, $P(A^c) = 1 - P(A)$
  2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, $P(A) \leq P(B)$ if $A \subseteq B$
  3. $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$
- A measurable space with probability measure is called the probability space

Basic combinatorics

- Factorial $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ (convention $0! = 1$)
- Binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
  - $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$
- Basic problem: an urn has n different balls. How many possible combinations of k balls can be drawn from the urn.
  1. Ordered with replacement n^k
  2. Ordered without replacement $n!/(n-k)!$
  3. Unordered with replacement $\binom{n+k-1}{k}$
  4. Unordered without replacement $\binom{n}{k}$

Independence and conditional probability

- Events A and B are independent iff $P(A \cap B) = P(A)\,P(B)$
- Let $(\Omega, \mathcal{F}, P)$ be the probability space and $A, B \in \mathcal{F}$. If $P(B) > 0$ then conditional probability of event A given event B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Let $A \in \mathcal{F}$ and $\{B_k\}_{k=1}^n$ such that $B_k \in \mathcal{F}$ and $B_i \cap B_j = \emptyset$ such that $A = \cup_{j=1}^n B_j$ then the *law of total probability* is

$$P(A) = \sum_{k=1}^{n} P(A|B_k)\,P(B_k).$$

- Suppose that $\{B_k\}_{k=1}^n$ satisfy conditions above, then if $P(A) > 0$ then the *Bayes' law* is

$$P(B_i \mid A) = \frac{P(A|B_i)P(B_i)}{\sum_{k=1}^n P(A|B_k)P(B_k)}$$

Random variables and distributions

- Recall that a real-valued random variable on $(\Omega, \mathcal{F}, P)$ (or, simply, a random variable) is a measurable function $\xi : \Omega \mapsto \mathbb{R}$, i.e. such that for any element of the Borel sigma algebra B, $\xi^{-1}(B) = \{\omega : \xi(\omega) \in B\}$ is an element of σ-algebra $\mathcal{F}$
- Notion of random variable allows us to abstract from (possibly complex) measurable space $(\Omega, \mathcal{F})$ and instead work with $(\mathbb{R}, \mathcal{B})$
- The *distribution* of random variable ξ is the probability measure $P_\xi(B) = P(\xi(\omega) \in B)$
- Random variable induces the probability space $(\mathbb{R}, \mathcal{B}, P_\xi)$
- Setting $B = (-\infty, x)$, we define the *distribution function* of random variable ξ as

$$F_\xi(x) = P(\xi < x)$$

- **Example: (Bernoulli r.v.)** $\Omega = \{0, 1\}$, "Probability of success" $p \in [0, 1]$. $P(\omega) = p^\omega (1-p)^{1-\omega}$. Random variable $X(\omega) = \omega$ for $\omega \in \Omega$.

- **Example: (binomial r.v.)** Sum of the sequence of k Bernoulli trials with probability of success p. $\Omega = \{\text{all sequences of } k \text{ zeros and ones}\}$. X takes integer values from 0 to k. Probability $P(X = r) = \binom{k}{r} p^r (1-p)^{k-r}$, $F(x) = \sum_{r<x} \binom{k}{r} p^r (1-p)^{k-r}$

- Distribution function $F(x)$ of r.v. X has the following properties

  1. Monotonicity: If $x_1 \leq x_2$ then $F(x_1) \leq F(x_2)$
  2. $\lim\limits_{x\to-\infty} F(x) = 0$ and $\lim\limits_{x\to+\infty} F(x) = 1$
  3. Left-continuity: $\lim\limits_{x\uparrow x_0} F(x) = F(x_0)$

- *(Kolmogorov's theorem)* If function $F(\cdot)$ satisfies conditions 1-3 above, then there exists a probability space $(\Omega, \mathcal{F}, P)$ and random variable X defined on this space such that $F_X(x) = F(x)$

- Common probability distributions

  1. Degenerate distribution: $F_a(x) = \mathbf{1}\{x > a\}$ ($\mathbf{1}\{\cdot\} \in \{0, 1\}$ is the indicator of $\cdot$), support on $\mathbb{R}$
  2. Binomial distribution: $F(x) = \sum_{r<x} \binom{k}{r} p^r (1-p)^{k-r}$, support on $\{0, 1, \ldots, k\}$
  3. Poisson distribution: $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$, support on $\{0, 1, \ldots, k, \ldots\}$
  4. Uniform distribution on $[a, b]$: $P(X \in B) = \int_{B\cap[a,b]} dx$
  5. Normal distribution: $P(X \in B) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_B e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$
  6. Exponential law: $P(X \in B) = \alpha \int_{B\cap(0,+\infty)} e^{-\alpha x} dx$

- The distribution of random variable X is *discrete* if X can take a countable number of values, such that $p_k = P(X = x_k)$ and $\sum p_k = 1$

- *Which listed distributions are discrete?*

- Distribution P of r.v. X is called *absolutely continuous* if for any Borel set B

$$P(B) = P(X \in B) = \int_B f(x)\, dx,$$

  where $f(x) \geq 0$ and $\int_{-\infty}^{+\infty} f(x)\, dx = 1$

- Function $f(\cdot)$ is called *the distribution density* of X.

- Suppose that $g(x)$ is a measurable function. Then if X is a r.v. then $Y = g(X)$ is also a random variable
$$F_{g(X)}(x) = P(g(X) < x) = P(X \in g^{-1}(-\infty, x))$$

  1. If $g(\cdot)$ is non-decreasing and its inverse is well-defined, then
  $$F_{g(X)}(x) = F_X(g^{-1}(x))$$
  2. If $g(\cdot)$ is strictly monotone and differentiable and X has a density $f(x)$ then
  $$f_{g(X)}(y) = f(g^{-1}(y)) \left|(g^{-1}(y))'\right|$$
  3. Useful analytic tool is *Leibnitz' rule:*
  $$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)\,dx = f(b(t),t)b'(t) - f(a(t),t)a'(t) + \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x,t)\,dx.$$
  4. **Example:** $X \sim U[0,1]$, $Y = -\log X$
  5. **Example:** $X \sim N(0,1)$, $Y = X^2$ (use cdf approach to get the answer, and recall that $|x| < a \Leftrightarrow -a < x < a$!)

Multivariate random variables and distributions

- Let $X_1, \ldots, X_n$ be r.v. defined on $(\Omega, \mathcal{F}, P)$. Mapping $\Omega \mapsto \mathbb{R}^n$ is called a *random vector* or a *multivariate random variable*

- $P_X(B)$ is the distribution of r.v. X (v. stands for vector). Function
$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P(X_1 < x_1, \ldots, X_n < x_n)$$
is the distribution function of r.v. X (or the joint distribution function of X)

- Joint distribution function maintains the properties 1-3 for one-dimensional case. In addition:

  1. $\lim\limits_{x_n \to +\infty} F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1})$

2. $\lim\limits_{x_n \to -\infty} F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = 0$

- As in one-dimensional case we can consider discrete and absolutely continuous distributions

- For absolutely continuous distributions we can provide equivalent definition of the cdf

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \int_{\infty}^{x_1} \cdots \int_{\infty}^{x_n} f(t_1,\ldots,t_n)\, dt_1 \ldots dt_n.$$

- Whenever the density exists, we can write an expression for it almost everywhere

$$\frac{\partial^n F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{\partial x_1 \ldots \partial x_n} = f(x_1,\ldots,x_n).$$

- If X has the density then any of its subvectors has a density. The density of a given subvector is called its *marginal density* and can be found as

$$f(x_1,\ldots,x_s) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1,\ldots,x_s,t_{s+1},\ldots,t_n)\, dt_n$$

- Suppose that in some $A \subset \mathbb{R}^n$ there are n continuously differentiable functions $y_i = g_i(x_1,\ldots,x_n)$ which are bijective in $x_1,\ldots,x_n$, i.e. there exist functions γ_i such that $x_i = \gamma_i(y_1,\ldots,y_n)$ and the Jacobian, defined as

$$J = \det \begin{pmatrix} \frac{\partial \gamma_1}{\partial y_1} & \cdots & \frac{\partial \gamma_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \gamma_n}{\partial y_1} & \cdots & \frac{\partial \gamma_n}{\partial y_n} \end{pmatrix} \neq 0 \text{ in } A.$$

  Then $Y_i = g_i(X_1,\ldots,X_n)$ are random variables with joint density

$$f_Y(y_1,\ldots,y_n) = f_X(x_1,\ldots,x_n)|J|$$

- **Example:** Consider joint distribution of $X_1 \sim U[0,2\pi]$ and $X_2 \sim$ Exponential(1) and consider random variables $Y_1 = X_2 \cos(X_1)$ and $Y_2 = X_2 \sin(X_1)$

- Random variables $X_1,\ldots,X_n$ are called *independent* if

$$P(X_1 \in B_1,\ldots,X_n \in B_n) = P(X_1 \in B_1)\ldots P(X_n \in B_n)$$

- **Theorem:** Random variables $X_1, \dots, X_n$ are independent iff

$$F_{X_1 \dots X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \dots F_{X_n}(x_n)$$

- **Corollary:** Suppose that r.v. $X_1, \dots, X_n$ have a joint density. Then they are independent iff

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

Characteristics of random variables

- The *expectation* of random variable X is the quantity

$$E[X] = \int_\Omega X(\omega) \, P(d\omega)$$

1. Equivalently $E[X] = \int x \, dF(x)$
2. If $F(x)$ has density $f(x)$ then $E[X] = \int x \, f(x) \, dx$
3. Expectation of X exists if $E[|X|] < \infty$ (this is violated if $1 - F(x) > 1/x$)
4. **Example:** Standard Cauchy distribution has density $f(x) = \frac{1}{\pi(1+x^2)}$. Verify that $E[X]$ does not exist noticing that $\frac{1}{2}(1+x^2)' = x$.

- Main properties of expectations

1. If a, b are constants then $E[a + bX] = a + bE[X]$
2. $E[X_1 + X_2] = E[X_1] + E[X_2]$ (if expectations exist)
3. If $a \le X \le b$ then $a \le E[X] \le b$; $E[X] \le E[|X|]$
4. If $X \ge 0$ and $E[X] = 0$ then $X \equiv 0$ with probability 1.
5. Probability can be defined via expectation

$$P(A) = E[\mathbf{1}\{A\}]$$

- Suppose that X, Y are r.v. and $g(x, y)$ is a measurable function.

1. Then if E[g(X,Y)] exists, then

$$E[g(X, Y)] = E[E[x, Y]_{x=X}]$$

2. If $g(x,y) = g_1(x)g_2(y)$, then

$$E[g(X,Y)] = E[g_1(X)]E[g_2(Y)]$$

- **Examples:**

1. Bernoulli r.v. $E[X] = 0 \cdot P(X=0) + 1 \cdot P(X=1) = p$
2. Binomial r.v. $X = X_1 + \ldots + X_k$ where $X_i \sim \text{Bernoulli}(p)$. Then $E[X] = k\,p$
3. Poisson r.v.

$$E[X] = \sum_{i=0}^{\infty} i \frac{\lambda^i}{i!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = \lambda,$$

given that $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$.

4. Normal r.v. with parameters (μ, σ^2)

$$E[X] = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\, dt$$

Change of variable $t = \frac{x-\mu}{\sigma}$:

$$E[X] = \mu \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt + \sigma \int_{-\infty}^{+\infty} t \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

first integral = 1 by definition of density, second integral = 0 given that integrand is an odd function.

5. Uniform $[a,b]$ r.v.: note for $U[0,1]$. $E[X] = \int_0^1 x\, dx = \frac{1}{2}$. $Y = a + (b-a)\,X \sim U[a,b]$. Thus $E[Y] = \frac{a+b}{2}$

Conditional expectations

- Let $(\Omega, \mathcal{F}, P)$ be the probability space and event $B \in \mathcal{F}$ is such that $P(B) > 0$. We can then form a new probability space $(\Omega, \mathcal{F}, P_B)$ where for each $A \in \mathcal{F}$

$$P_B(A) = P(A|B)$$

- We can verify that $(\Omega, \mathcal{F}, P_B)$ is indeed a probability space and any r.v. X in $(\Omega, \mathcal{F}, P)$ is also a r.v. in $(\Omega, \mathcal{F}, P_B)$

- The expectation of X in $(\Omega, \mathcal{F}, P_B)$ is called the *conditional expectation* of random variable X conditional on B:

$$E[X \mid B] = \int_\Omega X(\omega) P(d\omega)$$

- Function $F(x|B) = P_B(X < x) = P(X < x \mid B)$ is a distribution function of X on $(\Omega, \mathcal{F}, P_B)$. It is called the *conditional distribution function* of X

- **Example:** Suppose that $F_X(x) = 1 - e^{-\mu x}$. Find $F(x|X \geq a)$.

Variance of r.v.'s

- Variance of r.v. X

$$V[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

- Can define variance as a minimum

$$V[X] = \min_a E[(X - a)^2]$$

1. Verify by using properties of expectation
2. Conclude that mean is the most accurate approximation of the random variable in mean square

- $\sqrt{V[X]}$ is called the standard deviation

- **Examples**:

1. Bernoulli r.v. $V[X] = E[X^2] - E[X]^2$ and $X^2 = X$, thus $V[X] = p(1-p)$
2. Poisson r.v. $V[X] = E[X^2] - E[X]^2$, $E[X^2] = e^{-\lambda} \sum_{i=0}^\infty i^2 \frac{\lambda^i}{i!} = \lambda e^{-\lambda} \sum_{i=0}^\infty \frac{(\lambda^i)'}{(i-1)!} = \lambda e^{-\lambda} \left(\sum_{i=1}^\infty \frac{\lambda^i}{(i-1)!} \right)'$. Since $\sum_{i=1}^\infty \frac{\lambda^i}{(i-1)!} = \lambda\, e^\lambda$, then

$$E[X^2] = \lambda e^{-\lambda} \left(\lambda\, e^\lambda \right)' = \lambda + \lambda^2.$$

Thus $V[X] = \lambda$.

3. Normal r.v. $V[X] = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$. Change of variable $t = (x-\mu)/\sigma$, leading to

$$V[X] = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2\, e^{-t^2/2}\, dt = -\frac{\sigma^2}{\sqrt{2\pi}} t\, e^{-t^2/2} |_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2} = \sigma^2$$

- Selected properties of the variance
  - $V[X] \geq 0$ and $V[X] = 0$ iff $P(X(\omega) = c) = 1$ (c is a constant)
  - If a, b are constants then
$$V[a + b\,X] = b^2\,V[X]$$
  - If X, Y are independent, then
$$V[X + Y] = V[X] + V[Y]$$

Covariance, correlation and moments

- *Covariance* between r.v. X, Y is the number $\mathrm{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$
- *Correlation coefficient* between r.v. X, Y with $V[X], V[Y] \neq 0$ is the number $\rho(X, Y) = \mathrm{Cov}(X, Y)/\sqrt{(}V[X]\,V[Y])$
  1. $|\rho(X, Y)| \leq 1$
  2. If X, Y are independent then $\rho(X, Y) = 0$
  3. $|\rho(X, Y)| = 1$ iff there exist constants a, b such that $P(Y = a + bX) = 1$
  4. If $\rho(X, Y) > 0$, then X, Y are positively correlated
- The *k-th moment* of r.v. X is $E[X^k]$
- The *central k-th moment* of r.v. X is $E[(X - E[X])^k]$
- The *mixed k-th moment* of r.v. X is $E[X_1^{k_1} \cdot \ldots \cdot X_n^{k_n}]$ with $\sum_{i=1}^{n} k_i = k$

Inequalities

- Cauchy-Schwarz inequality: $E[|X\,Y|] \leq \sqrt{E[X^2]\,E[Y^2]}$
  - Obtain from $2|ab| \leq a^2 + b^2$ with $a = X^2/E[X^2]$ and $b = Y^2/E[Y^2]$
- Hölder's inequality: for $r > 1$ and $r^{-1} + s^{-1} = 1$ $E[|X\,Y|] \leq (E[X^r])^{1/r}\,(E[Y^s])^{1/s}$
  - Obtain from convexity of function x^r ($r > 1$) which leads to $r(x - 1) \leq x^r - 1$ and set $x = (a/b)^{1/r}$

- Jensen's inequality: for convex function $g(\cdot)$, $g(E[X]) \leq E[g(X)]$
  - Obtain from $g(x) \geq g(y) + (x-y)g'(y)$ and set $x = X$ and $y = E[X]$
- $P(|X| \geq \epsilon) \leq \frac{E[|X|]}{\epsilon}$
- Chebychev's inequality: $P(|X - E[X]| \geq \epsilon) \leq \frac{V[X]}{\epsilon^2}$

Characteristic functions

- *Complex numbers* are extension of real numbers
  1. Equation $x^k = -1$ does not have solutions for even k but has solutions for odd
  2. To resolve this asymmetry introduce the *complex domain*, denoted $\mathcal{C}$
  3. Basic construct is *imaginary unit* $i = \sqrt{-1}$
  4. Complex domain is constructed from numbers $z = x + i\, y$, where $x, y \in \mathbb{R}$ and $i = \sqrt{-1}$
  5. x is called the real part of z (denoted $\mathrm{Re}(z)$) and y is called the imaginary part of z (denoted $\mathrm{Im}(z)$)
  6. The absolute value of complex number is $\rho = |z| = \sqrt{x^2 + y^2}$
  7. The expression for complex numbers in polar form $z = \rho(\cos\phi + i\sin\phi)$, where $\phi = \tan^{-1}(y/x)$
  8. Euler's formula: $z = \rho\, e^{i\,\phi}$
  9. Using Euler's formula, we can easily solve equations like $x^k = -1$, since $\cos(\pi + 2\pi\, r) = -1$ for $r \in \{\ldots, -1, 0, 1, 2, \ldots\}$
- Using complex numbers also helps with analysis of random variables (recall issues with Cauchy distribution)
- The *characteristic function* of real-valued r.v. X is a complex-valued function

  $$\phi_X(t) = E[e^{itX}] = \int e^{itx}\, dF(x)$$

  c.f. exists for any r.v.

  1. For any r.v. X $\phi_X(0) = 1$ and $|\phi_X(t)| \leq 1$

2. $\phi_{aX+b}(t) = e^{itb}\phi_X(at)$

3. If $X_1, \dots, X_n$ are independent r.v. then

$$\phi_{X_1+\ldots+X_n}(t) = \phi_{X_1}(t)\ldots\phi_{X_n}(t)$$

4. $\phi_X(t)$ is uniformly continuous

5. If k-th moment of X exists, then $\phi_X(t)$ has a continuous k-th derivative and $(\phi_X(0))^{(k)} = i^k\, E[X^k]$

- **Examples:**

1. Degenerate r.v. $X = a$ w.p. 1, then $\phi_X(t) = e^{ita}$

2. Bernoulli (p) r.v. $\phi_X(t) = p(e^{it} - 1) + 1$

3. Normal $(0, 1)$ r.v.

$$\phi_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}}\, dx$$

Differentiate by t and integrate by parts

$$\phi_X'(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ixe^{itx-\frac{x^2}{2}}\, dx = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} te^{itx-\frac{x^2}{2}}\, dx = -t\,\phi_X(t)$$

This means that

$$(\log\, \phi_X(t))' = -t, \;\; \log\, \phi_X(t) = -\frac{t^2}{2} + c$$

$$\phi_X(t) = e^{-t^2/2}$$

4. Poisson (λ) r.v.

$$\phi_X(t) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} \frac{(\lambda\, e^{it})^k}{k!} e^{-\lambda} = e^{-\lambda}\, e^{\lambda\, e^{it}} = e^{\lambda(e^{it}-1)}$$

5. Standard Cauchy r.v. $\phi_X(t) = e^{-|t|}$

- **Theorem** Continuous function $\phi(t)$ with $\phi(0) = 1$ is a characteristic function iff it is positive semi-definite, i.e. for any $t_1, \dots, t_n \in \mathbb{R}$ and any $\lambda_1, \dots \lambda_n \in \mathbb{C}$

$$\sum_{k,j=1}^{n} \phi(t_k - t_j)\lambda_k\lambda_j^* \geq 0$$

(λ^* is the complex conjugate)

- Sufficiency is clear, since if $\phi(t) = E[e^{itX}]$, then

$$\sum_{k,j=1}^{n} \phi(t_k - t_j)\lambda_k \lambda_j^* = E[\sum_{k,j=1}^{n} e^{i(t_k - t_j)X} \lambda_k \lambda_j^*] = E\left[\left|\sum_{k=1}^{n} \lambda_k e^{it_k X}\right|^2\right] \geq 0$$

- **Theorem** If X has a density then $\lim\limits_{t \to \infty} \phi_X(t) = 0$
- Intuition for the proof: If $f(x)$ has a derivative or order k, then

$$\phi_X(t) = \int e^{itx} f(x)\, dx = \frac{1}{it} \int e^{itx} f'(x)\, dx = \ldots = \frac{1}{(it)^k} \int e^{itx} f^{(k)}(x)\, dx.$$

  Thus $|\phi_X(t)| \leq \frac{c}{|t|^k}$
- **Theorem:** The characteristic function of X uniquely defines its distribution.

Moment-generating functions

- The *moment-generating function* of r.v. X is a real-valued function

$$M_X(t) = E[e^{tX}]$$

  1. It is not guaranteed to exist
  2. If k-th moment of r.v. X exists then $(M_X(0))^{(k)} = E[X^k]$
  3. $M_{aX+b}(t) = e^{tb} M_X(at)$
  4. If $X_1, \ldots, X_n$ are independent r.v.and their m.g.f. exist then

$$M_{X_1 + \ldots + X_n}(t) = M_{X_1}(t) \ldots M_{X_n}(t)$$