Probability Theory

Lecture 3

Per Sidén

Division of Statistics
Dept. of Computer and Information Science
Linköping University

Overview Lecture 3

- Transforms
- Probability generating function
- Moment generating function
- Characteristic function
- Transforms and distributions with random parameters

TRANSFORMS

- Finding the distribution of sum of random variables is hard. Convolution is messy.

Transforms

- Finding the distribution of sum of random variables is hard. Convolution is messy.
- Transforms are functions that *uniquely* describe probability distributions.

Transforms

- Finding the distribution of sum of random variables is hard. Convolution is messy.
- Transforms are functions that *uniquely* describe probability distributions.
- If you know the transform, you know the distribution, and vice versa.

Transforms

- Finding the distribution of sum of random variables is hard. Convolution is messy.
- Transforms are functions that *uniquely* describe probability distributions.
- If you know the transform, you know the distribution, and vice versa.
- $X \stackrel{d}{=} Y \iff g_X(t) = g_Y(t)$

Transforms

- Finding the distribution of sum of random variables is hard. Convolution is messy.
- Transforms are functions that *uniquely* describe probability distributions.
- If you know the transform, you know the distribution, and vice versa.
- $X \stackrel{d}{=} Y \iff g_X(t) = g_Y(t)$
- **Summation** of independent variables corresponds to **multiplication of transforms**. Nice!

Probability generating function

- Applies to **non-negative**, **integer-valued** random variables.

Probability generating function

- Applies to **non-negative**, **integer-valued** random variables.

DEF The **probability generating function** of X is

$$g_X(t) = \mathrm{E}t^X = \sum_{n=0}^{\infty} t^n \cdot P(X = n)$$

Probability generating function

- Applies to **non-negative**, **integer-valued** random variables.

DEF The **probability generating function** of X is

$$g_X(t) = \mathrm{E}t^X = \sum_{n=0}^{\infty} t^n \cdot P(X = n)$$

- $g_X(t)$ is defined at least for $|t| \leq 1$.

Probability generating function

- Applies to **non-negative**, **integer-valued** random variables.

DEF The **probability generating function** of X is

$$g_X(t) = \mathrm{E}t^X = \sum_{n=0}^{\infty} t^n \cdot P(X = n)$$

- $g_X(t)$ is defined at least for $|t| \leq 1$.

TH If $g_X = g_Y$ then $p_X = p_Y$.

Probability generating function

- Applies to **non-negative**, **integer-valued** random variables.

DEF The **probability generating function** of X is

$$g_X(t) = \mathrm{E}t^X = \sum_{n=0}^{\infty} t^n \cdot P(X = n)$$

- $g_X(t)$ is defined at least for $|t| \leq 1$.

TH If $g_X = g_Y$ then $p_X = p_Y$.

-

TH Let $X_1, X_2, ..., X_n$ be independent. Then

$$g_{X_1+X_2+...+X_n}(t) = \prod_{k=1}^{n} g_{X_k}(t)$$

Probability generating function, cont.

Cor Let $X_1, X_2, ..., X_n$ be independent and identically distributed. Then

$$g_{X_1+X_2+...+X_n}(t) = (g_X(t))^n$$

Probability generating function, cont.

Cor Let $X_1, X_2, ..., X_n$ be independent and identically distributed. Then

$$g_{X_1+X_2+...+X_n}(t) = (g_X(t))^n$$

- The name probability generating function comes from:

$$P(X = n) = \frac{g_X^{(n)}(0)}{n!}$$

where $g_X^{(n)}(t)$ is the nth derivative of $g_X(t)$ wrt to t.

PROBABILITY GENERATING FUNCTION, CONT.

COR Let $X_1, X_2, ..., X_n$ be independent and identically distributed. Then

$$g_{X_1+X_2+...+X_n}(t) = (g_X(t))^n$$

- The name probability generating function comes from:

$$P(X = n) = \frac{g_X^{(n)}(0)}{n!}$$

where $g_X^{(n)}(t)$ is the nth derivative of $g_X(t)$ wrt to t.

TH Factorial moments (if $\mathrm{E}|X|^k < \infty$)

$$\mathrm{E}(X(X-1)\cdots(X-k+1)) = g_X^{(k)}(1)$$

Probability generating function, cont.

Cor Let $X_1, X_2, ..., X_n$ be independent and identically distributed. Then

$$g_{X_1+X_2+...+X_n}(t) = (g_X(t))^n$$

- The name probability generating function comes from:

$$P(X = n) = \frac{g_X^{(n)}(0)}{n!}$$

where $g_X^{(n)}(t)$ is the nth derivative of $g_X(t)$ wrt to t.

Th Factorial moments (if $\mathrm{E}\,|X|^k < \infty$)

$$\mathrm{E}\left(X(X-1)\cdots(X-k+1)\right) = g_X^{(k)}(1)$$

- Moments can be computed

$$\mathrm{E}X = g_X'(1)$$

$$\mathrm{Var}X = g_X''(1) + g_X'(1) - \left(g_X'(1)\right)^2$$

Probability generating function - examples

- ✓ binomial theorem: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.
- ✏ Bernoulli, $X \sim Be(p)$

Probability generating function - examples

- ✓ binomial theorem: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.
- ✏ Bernoulli, $X \sim Be(p)$

$$g_X(t) = \sum_{n=0}^{\infty} t^n \cdot P(X=n) = t^0 q + t^1 p = q + pt$$

- ✏ Binomial, $X \sim Bin(n,p)$

Probability generating function - examples

- ✓ binomial theorem: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.
- ✏ Bernoulli, $X \sim Be(p)$

$$g_X(t) = \sum_{n=0}^{\infty} t^n \cdot P(X=n) = t^0 q + t^1 p = q + pt$$

- ✏ Binomial, $X \sim Bin(n,p)$

$$g_X(t) = \sum_{k=0}^{n} t^k \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pt)^k q^{n-k} = (q+pt)^n$$

- ✏ Let $X_1, ..., X_n \overset{iid}{\sim} Be(p)$, then what is $X = X_1 + ... + X_n$?

PROBABILITY GENERATING FUNCTION - EXAMPLES

- ✓ binomial theorem: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.
- ✏ Bernoulli, $X \sim Be(p)$

$$g_X(t) = \sum_{n=0}^{\infty} t^n \cdot P(X=n) = t^0 q + t^1 p = q + pt$$

- ✏ Binomial, $X \sim Bin(n,p)$

$$g_X(t) = \sum_{k=0}^{n} t^k \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pt)^k q^{n-k} = (q+pt)^n$$

- ✏ Let $X_1, ..., X_n \overset{iid}{\sim} Be(p)$, then what is $X = X_1 + ... + X_n$?

$$g_X(t) = \prod_{i=1}^{n} g_{X_i(t)} = \prod_{i=1}^{n} (q+pt) = (q+pt)^n$$

so $X \sim Bin(n,p)$.

PROBABILITY GENERATING FUNCTION - EXAMPLES

- ✓ Poisson prob func: $p(X = k) = e^{-m}m^k/k!$
- ✓ $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$
- ✏ Poisson, $X \sim Po(m)$

Probability generating function - examples

- ✓ Poisson prob func: $p(X = k) = e^{-m} m^k / k!$
- ✓ $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$
- ✏ Poisson, $X \sim Po(m)$

$$g_X(t) = \sum_{k=0}^{\infty} t^k \frac{e^{-m} m^k}{k!} = e^{-m} \sum_{k=0}^{\infty} \frac{(mt)^k}{k!} = e^{m(t-1)}$$

- ✏ If $X_1 \sim Po(m_1)$ independently of $X_2 \sim Po(m_2)$, what is $X_1 + X_2$?

Probability generating function - examples

- ✓ Poisson prob func: $p(X = k) = e^{-m}m^k/k!$
- ✓ $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$
- ✏ Poisson, $X \sim Po(m)$

$$g_X(t) = \sum_{k=0}^{\infty} t^k \frac{e^{-m}m^k}{k!} = e^{-m} \sum_{k=0}^{\infty} \frac{(mt)^k}{k!} = e^{m(t-1)}$$

- ✏ If $X_1 \sim Po(m_1)$ independently of $X_2 \sim Po(m_2)$, what is $X_1 + X_2$?

$$g_{X_1+X_2}(t) = e^{m_1(t-1)}e^{m_2(t-1)} = e^{(m_1+m_2)(t-1)}$$

so $X_1 + X_2 \sim Po(m_1 + m_2)$.

Moment generating function

- $g_X(t)$ limited to non-negative integer-valued variables.

Moment generating function

- $g_X(t)$ limited to non-negative integer-valued variables.

DEF **Moment generating function** of a variable X

$$\psi_X(t) = \mathrm{E}e^{tX}$$

if the expectation exist and is finite for $|t| < h$, for some $h > 0$.

Moment generating function

- $g_X(t)$ limited to non-negative integer-valued variables.

DEF **Moment generating function** of a variable X

$$\psi_X(t) = \mathrm{E}e^{tX}$$

if the expectation exist and is finite for $|t| < h$, for some $h > 0$.

TH If $\psi_X(t)$ exists for $|t| < h$ for some $h > 0$, then

- All moments exist $\mathrm{E}|X|^r < \infty$ for all $r > 0$
- $\mathrm{E}X^n = \psi_X^{(n)}(0)$ for $n = 1, 2, ...$

Moment generating function

- $g_X(t)$ limited to non-negative integer-valued variables.

Def **Moment generating function** of a variable X

$$\psi_X(t) = \mathrm{E}e^{tX}$$

if the expectation exist and is finite for $|t| < h$, for some $h > 0$.

Th If $\psi_X(t)$ exists for $|t| < h$ for some $h > 0$, then

  - All moments exist $\mathrm{E}|X|^r < \infty$ for all $r > 0$
  - $\mathrm{E}X^n = \psi_X^{(n)}(0)$ for $n = 1, 2, ...$

- Taylor expansion around $t = 0$ [note $\frac{\partial^k e^{tX}}{\partial t^k} = X^k e^{tX}$]

$$e^{tX} = 1 + \sum_{n=1}^{\infty} \frac{t^n X^n}{n!}$$

so

$$Ee^{tX} = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} EX^n$$

Moment generating function - examples

- $X \sim Be(p)$

Moment generating function - examples

- $X \sim Be(p)$

$$\psi_X(t) = Ee^{tX} = qe^{t \cdot 0} + pe^{t \cdot 1} = q + pe^t$$

  - $\psi_X'(t) = pe^t$ so $E(X) = \psi_X'(0) = p$.
  - $\psi_X''(t) = pe^t$ so $E(X^2) = \psi_X''(0) = p$.
  - $Var(X) = E(X^2) - [E(X)]^2 = p - p^2 = pq$

- $X \sim \Gamma(p, a)$

Moment generating function - examples

- $X \sim Be(p)$

$$\psi_X(t) = Ee^{tX} = qe^{t\cdot 0} + pe^{t\cdot 1} = q + pe^t$$

  - $\psi'_X(t) = pe^t$ so $E(X) = \psi'_X(0) = p$.
  - $\psi''_X(t) = pe^t$ so $E(X^2) = \psi''_X(0) = p$.
  - $Var(X) = E(X^2) - [E(X)]^2 = p - p^2 = pq$

- $X \sim \Gamma(p, a)$

$$\psi_X(t) = \frac{1}{(1-at)^p}$$

  - $\psi'_X(t) = \frac{ap}{(1-at)^{p+1}}$ so $E(X) = \psi'_X(0) = ap$.
  - $\psi''_X(t) = \frac{a^2 p(p+1)}{(1-at)^{p+2}}$ so $E(X^2) = \psi''_X(0) = a^2 p(p+1)$.
  - $Var(X) = E(X^2) - [E(X)]^2 = a^2 p(p+1) - a^2 p^2 = a^2 p$.

Moment generating function, cont.

Th If $\exists\ h > 0$ such that $\psi_X(t) = \psi_Y(t)$ for $|t| < h$, then $X \overset{d}{=} Y$.

Moment generating function, cont.

Th If $\exists\ h > 0$ such that $\psi_X(t) = \psi_Y(t)$ for $|t| < h$, then $X \overset{d}{=} Y$.

Th If $X_1, X_2, ..., X_n$ are independent with moment generating functions that exist for $|t| < h$ for some $h > 0$, then

$$\psi_{X_1 + \ldots X_n}(t) = \prod_{i=1}^{n} \psi_{X_i}(t), \quad t < |h|$$

Moment generating function, cont.

TH If $\exists\ h > 0$ such that $\psi_X(t) = \psi_Y(t)$ for $|t| < h$, then $X \stackrel{d}{=} Y$.

TH If $X_1, X_2, ..., X_n$ are independent with moment generating functions that exist for $|t| < h$ for some $h > 0$, then

$$\psi_{X_1+\ldots X_n}(t) = \prod_{i=1}^{n} \psi_{X_i}(t), \quad t < |h|$$

TH Moment generating function of a linear combination $a \cdot X + b$

$$\psi_{aX+b}(t) = e^{tb}\psi_X(at)$$

Moment generating function, cont.

TH If $\exists\ h > 0$ such that $\psi_X(t) = \psi_Y(t)$ for $|t| < h$, then $X \overset{d}{=} Y$.

TH If $X_1, X_2, ..., X_n$ are independent with moment generating functions that exist for $|t| < h$ for some $h > 0$, then

$$\psi_{X_1 + \ldots X_n}(t) = \prod_{i=1}^{n} \psi_{X_i}(t), \quad t < |h|$$

TH Moment generating function of a linear combination $a \cdot X + b$

$$\psi_{aX+b}(t) = e^{tb} \psi_X(at)$$

✏ If $X \sim \Gamma(d, p)$, what is the distribution of $Y = \sigma \cdot X$?

Moment generating function, cont.

Th If $\exists\ h > 0$ such that $\psi_X(t) = \psi_Y(t)$ for $|t| < h$, then $X \overset{d}{=} Y$.

Th If $X_1, X_2, ..., X_n$ are independent with moment generating functions that exist for $|t| < h$ for some $h > 0$, then

$$\psi_{X_1+\ldots X_n}(t) = \prod_{i=1}^{n} \psi_{X_i}(t), \quad t < |h|$$

Th Moment generating function of a linear combination $a \cdot X + b$

$$\psi_{aX+b}(t) = e^{tb}\psi_X(at)$$

✏ If $X \sim \Gamma(d, p)$, what is the distribution of $Y = \sigma \cdot X$?

$$\psi_X(t) = \frac{1}{(1-dt)^p}$$

$$\psi_Y(t) = \frac{1}{(1-d\sigma t)^p},$$

which is the mgf of $\Gamma(d\sigma, p)$. Gamma family is closed under scaling.

The characteristic function

- Moment generating function is not defined for all random variable. No mgf for Cauchy or LogNormal.

The characteristic function

- Moment generating function is not defined for all random variable. No mgf for Cauchy or LogNormal.
- The **characteristic function** is more general and exists for any variable, **but complex valued**.

The characteristic function

- Moment generating function is not defined for all random variable. No mgf for Cauchy or LogNormal.
- The **characteristic function** is more general and exists for any variable, **but complex valued**.

Def The characteristic function of a random variable X is

$$\varphi_X(t) = Ee^{itX} = E\left(\cos tX + i \sin tX\right)$$

where i is the imaginary number ($i^2 = -1$).

✏ $X \sim U(a, b)$

The characteristic function

- Moment generating function is not defined for all random variable. No mgf for Cauchy or LogNormal.
- The **characteristic function** is more general and exists for any variable, **but complex valued**.

DEF The characteristic function of a random variable X is

$$\varphi_X(t) = Ee^{itX} = E\left(\cos tX + i \sin tX\right)$$

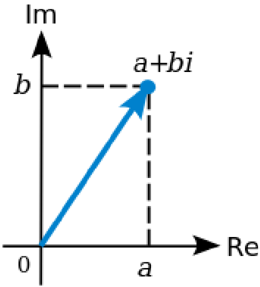
where i is the imaginary number ($i^2 = -1$).

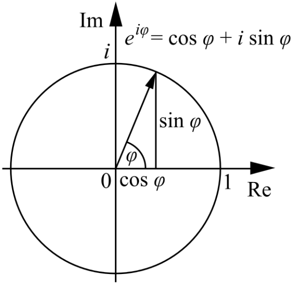
✏ $X \sim U(a, b)$, then

$$\varphi_X\left(t\right) = \frac{e^{itb} - e^{ita}}{it\left(b - a\right)}$$

Complex numbers

- Complex number $z = a + b \cdot i$
- $Re(z) = a$ is the real part of z
- $Im(z) = b$ is the imaginary part of z
- Complex conjugate $\bar{z} = a - b \cdot i$
- Addition: $z_1 + z_2 = a_1 + a_2 + (b_1 + b_2) \cdot i$
- Multiplication: $z_1 z_2 = a_1 a_2 - b_1 b_2 + (a_1 b_2 + a_2 b_1) i$
- Modulus: $|z| = \sqrt{a^2 + b^2}$. Length of vector.
- Complex exponentials: $e^{ix} = \cos x + i \cdot \sin x$





The characteristic function, cont.

Th If $\varphi_X = \varphi_Y$ then $X \overset{d}{=} Y$.

The characteristic function, cont.

TH If $\varphi_X = \varphi_Y$ then $X \stackrel{d}{=} Y$.

TH Let F be the distribution function of X. If F is continuous at a and b, and $\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty$ then

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

The characteristic function, cont.

Th If $\varphi_X = \varphi_Y$ then $X \stackrel{d}{=} Y$.

Th Let F be the distribution function of X. If F is continuous at a and b, and $\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty$ then

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

Th Characteristic function of a sums of independent variables

$$\varphi_{X_1 + \ldots + X_n}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t)$$

The characteristic function, cont.

Th If $\varphi_X = \varphi_Y$ then $X \overset{d}{=} Y$.

Th Let F be the distribution function of X. If F is continuous at a and b, and $\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty$ then

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

Th Characteristic function of a sums of independent variables

$$\varphi_{X_1 + \ldots + X_n}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t)$$

Th Moments

$$\varphi_X^{(k)}(0) = i^k \cdot EX^k$$

THE CHARACTERISTIC FUNCTION, CONT.

TH If $\varphi_X = \varphi_Y$ then $X \stackrel{d}{=} Y$.

TH Let F be the distribution function of X. If F is continuous at a and b, and $\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty$ then

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

TH Characteristic function of a sums of independent variables

$$\varphi_{X_1+\ldots+X_n}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t)$$

TH Moments

$$\varphi_X^{(k)}(0) = i^k \cdot EX^k$$

TH Linear combinations

$$\varphi_{aX+b}(t) = e^{ibt} \varphi_X(at)$$

Transforms - distributions with random parameters

- Transforms are expected values (or t^X, e^{tX} or e^{itX}), so the law of iterated expectation is useful.
- Let $X|(N = n) \sim Bin(n, p)$ and $N \sim Po(\lambda)$. What is the marginal distribution of X?

TRANSFORMS - DISTRIBUTIONS WITH RANDOM PARAMETERS

- Transforms are expected values (or t^X, e^{tX} or e^{itX}), so the law of iterated expectation is useful.
- Let $X|(N = n) \sim Bin(n, p)$ and $N \sim Po(\lambda)$. What is the marginal distribution of X? X is non-negative and integer-valued, so $g_X(t)$ is defined.

$$g_X(t) = E\left(E(t^X|N)\right) = Eh(N)$$

where

$$h(n) = E(t^X|N = n) = (q + pt)^n.$$

We then have

$$g_X(t) = E\left((q + pt)^N\right) = g_N(q + pt) = e^{\lambda[(q+pt)-1]} = e^{\lambda p(t-1)}.$$

- $X|y \sim N(0, y)$ and $y \sim \mathrm{Exp}(1)$, then $X \sim L(1/\sqrt{2})$. Prove using characteristic functions.

Transforms - sums of random number of random variables

Th Let $S_n = X_1 + X_2 + ... + X_n$ be a sum of i.i.d variables and N be a non-negative integer valued random variable. Then

$$\begin{aligned} g_{S_N}(t) &= g_N(g_X(t)) \\ \psi_{S_N}(t) &= g_N(\psi_X(t)). \\ \varphi_{S_N}(t) &= g_N(\varphi_X(t)) \end{aligned}$$

Transforms - sums of random number of random variables

Th Let $S_n = X_1 + X_2 + ... + X_n$ be a sum of i.i.d variables and N be a non-negative integer valued random variable. Then

$$\begin{aligned} g_{S_N}(t) &= g_N(g_X(t)) \\ \psi_{S_N}(t) &= g_N(\psi_X(t)). \\ \varphi_{S_N}(t) &= g_N(\varphi_X(t)) \end{aligned}$$

✏ $X_1, X_2, \ldots \sim Exp(1)$ (i.i.d) and $N \sim Fs(p)$. S_N?

Transforms - sums of random number of random variables

Th Let $S_n = X_1 + X_2 + ... + X_n$ be a sum of i.i.d variables and N be a non-negative integer valued random variable. Then

$$\begin{aligned} g_{S_N}(t) &= g_N(g_X(t)) \\ \psi_{S_N}(t) &= g_N(\psi_X(t)). \\ \varphi_{S_N}(t) &= g_N(\varphi_X(t)) \end{aligned}$$

✏ $X_1, X_2, \ldots \sim Exp(1)$ (i.i.d) and $N \sim Fs(p)$. S_N?

$$\begin{aligned} \psi_{S_N}(t) &= g_N(\psi_X(t)) = \frac{1}{1-\frac{t}{p}} \\ &\Rightarrow S_N \sim Exp(1/p) \end{aligned}$$