3. Functions of random variables

- linear and quadratic transformations
- general transformations
- characteristic function
- Markov and Chebyshev inequalities
- Chernoff bound

Functions of random variables

let X be an RV and $g(x)$ be a real-valued function defined on the real line

- $Y = g(X)$, Y is also an RV
- CDF of Y will depend on $g(x)$ and CDF of X

Example: define $g(x)$ as

$$g(x) = (x)^+ = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

- an input voltage X passes thru a halfwave rectifier
- A/D converter: a uniform quantizer maps input to the closet point
- Y is # of active speakers in excess of M, *i.e.*, $Y = (X - M)^+$

CDF of $Y = g(X)$

probability of equivalent events:

$$P(Y \text{ in } C) = P(g(X) \text{ in } C) = P(X \text{ in } B)$$

where B is the equivalent event of values of X sucht that $g(X)$ is in C

Example: Voice Transmission System

- X is # of active speakers in a group of N speakers
- let p be the probability that a speaker is active
- a voice transmission system can transmit up to M signals at a time
- let Y be the number of signal discarded, so $Y = (X - M)^+$

Y take values from the set $S_Y = \{0, 1, \dots, N - M\}$

we can compute PMF of Y as

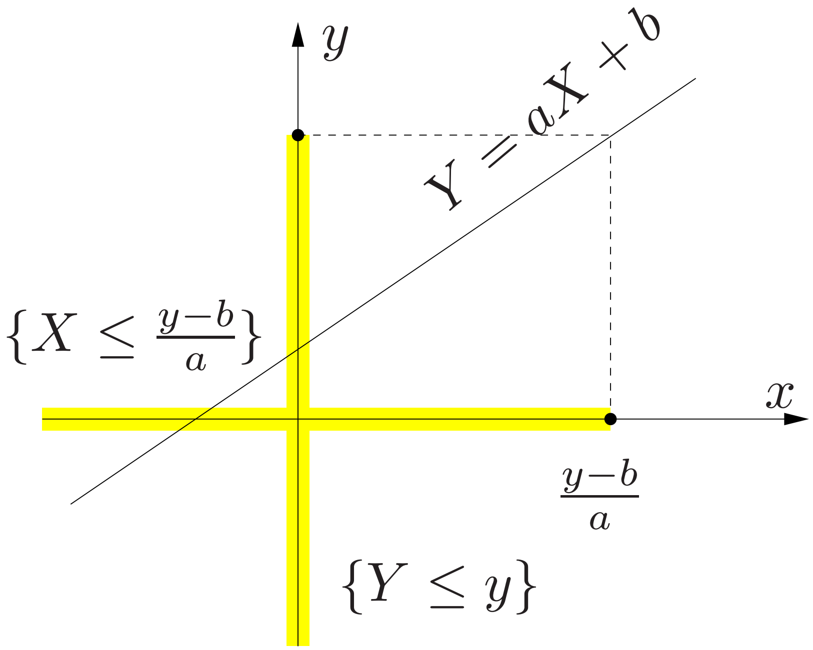
$$P(Y = 0) = P(X \text{ in } \{0, 1, \dots, M\}) = \sum_{k=0}^{M} p_X(k)$$

$$P(Y = k) = P(X = M + k) = p_X(M + k), \quad 0 < k \le N - M,$$

Affine functions

define $Y = aX + b, \quad a > 0$. Find CDF and PDF of Y

If $a > 0$



$$\begin{aligned} P(Y \leq y) &= P(aX + b \leq y) \\ &= P(X \leq (y-b)/a) \end{aligned}$$

thus,

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right)$$

PDF of Y is obtained by differentiating the CDF wrt. to y

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

Example: Affine function of a Gaussian

let $X \sim \mathcal{N}(m, \sigma^2)$:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{(x-m)^2}{2\sigma^2}$$

let $Y = aX + b$, with $a > 0$

from page 3-5,

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right) = \frac{1}{\sqrt{2\pi(a\sigma)^2}} \exp -\frac{(y-b-am)^2}{2(a\sigma)^2}$$

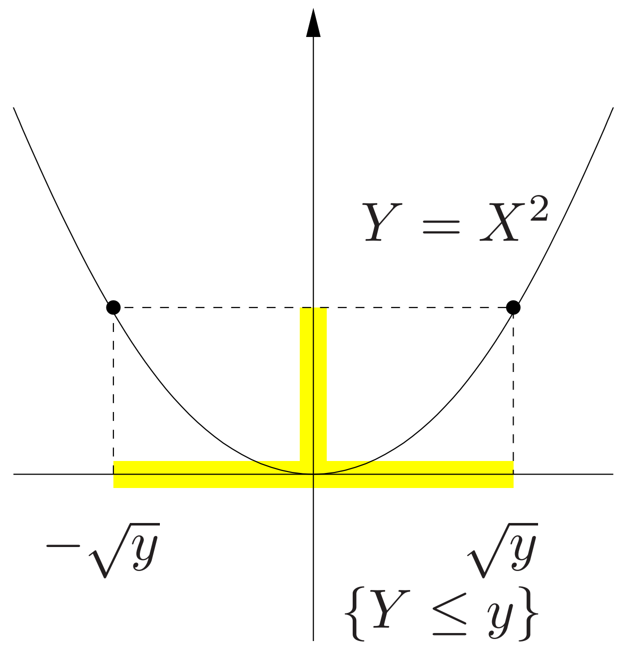
- Y has also a Gaussian distribution with mean $b + am$ and variance $(a\sigma)^2$
- thus, a linear function of a Gaussian is also a Gaussian

Example: Quadratic functions

define $Y = X^2$. find CDF and PDF of Y

for a positive y, we have

$$\{Y \leq y\} \Longleftrightarrow \{-\sqrt{y} \leq X \leq \sqrt{y}\}$$



thus,

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ F_X(\sqrt{y}) - F_X(-\sqrt{y}), & y > 0 \end{cases}$$

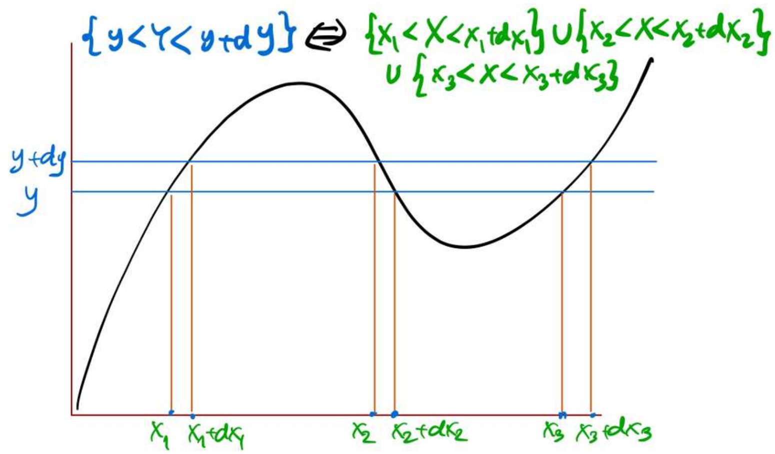
differentiating wrt. to y gives

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}$$

for $X \sim \mathcal{N}(0,1)$, Y is a Chi-square random variable with one DOF

General functions of random variables

suppose $Y = g(X)$ is a transformation (could be many-to-one)



suppose $y = g(x)$ has n roots:

$$y = g(x_1) = g(x_2) = \cdots = g(x_n)$$

two equivalent events: $\{y < Y < y + dy\} \iff \bigcup_{k=1}^{n} \{x_k < X < x_k + dx_k\}$

the probabibilities of two equivalent events are approximately

$$f_Y(y)|dy| = f_X(x_1)|dx_1| + f_X(x_2)|dx_2| + \cdots + f_X(x_n)|dx_n|$$

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \cdots + \frac{f_X(x_n)}{|g'(x_n)|}$$

where $g'(x)$ is the derivative (Jacobian) of $g(x)$

Affine transformation: $Y = aX + b$, $g'(x) = a$

the equation $y = ax + b$ has a single solution $x = (y - b)/a$ for every y, so

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y - b}{a}\right)$$

Quadratic transformation: $Y = aX^2$, $a > 0$, $g'(x) = 2ax$

if $y \leq 0$, then the equation $y = ax^2$ has no real solutions, so $f_Y(y) = 0$

if $y > 0$, then it has two solutions

$$x_1 = \sqrt{y/a}, \quad x_2 = -\sqrt{y/a}$$

and therefore

$$f_Y(y) = \frac{1}{2\sqrt{ay}}\Big(f_X(\sqrt{y/a}) + f_X(-\sqrt{y/a})\Big)$$

Log of uniform variables

verify that if X has a standard uniform distribution $\mathcal{U}(0,1)$, then

$$Y = -\log(X)/\lambda$$

has an exponential distribution with parameter λ

for $Y = y$, we can solve $X = x = e^{-\lambda y} \Rightarrow$ unique root

- the Jacobian is $g'(x) = -\frac{1}{\lambda x} = -e^{\lambda y}/\lambda$
- when $y < 0$, $x = e^{-\lambda y} \notin [0,1]$; hence, $f_Y(y) = 0$
- when $y \geq 0$ (or $e^{-\lambda y} \in [0,1]$), we will have

$$f_Y(y) = \frac{f_X(e^{-\lambda y})}{|-1/\lambda x|} = \lambda e^{-\lambda y}$$

Amplitude samples of a sinusoidal waveform

let $Y = \cos X$ where $X \sim \mathcal{U}(0, 2\pi]$, find the pdf of Y

for $|y| > 1$ there is no solution of $x \Rightarrow f_Y(y) = 0$

for $|y| < 1$ the equation $y = \cos x$ has two solutions:

$$x_1 = \cos^{-1}(y), \quad x_2 = 2\pi - x_1$$

the Jacobians are

$$g'(x_1) = -\sin(x_1) = -\sin(\cos^{-1}(y)) = -\sqrt{1-y^2}, \quad g'(x_2) = \sqrt{1-y^2}$$

since $f_X(x) = 1/2\pi$ in the interval $(0, 2\pi]$, so

$$f_Y(y) = \frac{1}{\pi\sqrt{1-y^2}}, \quad \text{for} \quad -1 < y < 1$$

note that although $f_Y(\pm 1) = \infty$ the probability that $y = \pm 1$ is 0

Transform Methods

- moment generating function
- characteristic function

Moment generating functions

for a random variable X, the moment generating function (MGF) of X is

$$\Phi(t) = \mathbf{E}[e^{tX}]$$

Continuous

$$\Phi(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Discrete

$$\Phi(t) = \sum_k e^{tx_k} p(x_k)$$

- except for a sign change, $\Phi(t)$ is the 2-sided Laplace transform of pdf
- the set of t for which the integral is finite forms the domain of $\Phi(t)$

Moment theorem

computing any moments of X is easily obtained by

$$\mathbf{E}[X^n] = \left.\frac{d^n\Phi(t)}{dt^n}\right|_{t=0}$$

because

$$\begin{aligned}\mathbf{E}[e^{tX}] &= \mathbf{E}\left[1 + tX + \frac{(tX)^2}{2!} + \cdots + \frac{(tX)^n}{n!} + \cdots\right] \\ &= 1 + t\mathbf{E}[X] + \frac{t^2}{2!}\mathbf{E}[X^2] + \cdots + \frac{t^n}{n!}\mathbf{E}[X^n] + \cdots\end{aligned}$$

note that $\Phi(0) = 1$

linear transformation: if $Y = aX + b$, then

$$\Phi_y(t) = e^{tb}\Phi_x(at)$$

MGF of Gaussian variables

the MGF of $X \sim \mathcal{N}(0,1)$ is given by

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

it can be derived by completing square in the exponent:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} e^{tx} dx = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx$$

the MGF of $X \sim \mathcal{N}(\mu, \sigma^2)$ (affine transformation of $\mathcal{N}(0,1)$) is

$$\Phi(t) = e^{(\mu t + \sigma^2 t^2/2)}$$

from the moment theorem, we obtain

$$\Phi'(0) = \mu, \quad \Phi''(0) = \mu^2 + \sigma^2$$

Characteristic functions

the characteristic function (CF) of a random variable X is defined by

Continuous

$$\Phi(\omega) = \mathbf{E}[e^{\mathrm{i}\omega X}] = \int_{-\infty}^{\infty} f(x) e^{\mathrm{i}\omega x} dx$$

Discrete

$$\Phi(\omega) = \mathbf{E}[e^{\mathrm{i}\omega X}] = \sum_{k} e^{\mathrm{i}\omega x_k} p(x_k)$$

- $\Phi(\omega)$ is simply the (inverse) Fourier transform of the PDF or PMF of X
- every pdf and its characteristic function form a unique Fourier pair:

$$\Phi(\omega) \Longleftrightarrow f(x)$$

- it looks as if we can obtain $\Phi(\omega)$ by substituting $t = \mathrm{i}\omega$ from MGF to CF but the existence of two transformations could be different

Properties of characteristic functions

- CF always exists because of absolute convergence (not true for MGF)

$$|\Phi(\omega)| \leq \int_{-\infty}^{\infty} |e^{\mathrm{i}\omega x}||f(x)|dx \int_{-\infty}^{\infty} f(x)dx = 1$$

- CF is maximum at origin because $f(x) \geq 0$:

$$|\Phi(\omega)| \leq \Phi(0) = 1$$

- CF is self-adjoint: $\Phi(-\omega) = \Phi^*(\omega)$ (where $*$ is complex conjugate)
- CF is non-negative definite: for any real numbers $w_1, w_2, \ldots, w_n$ and complex numbers $z_1, z_2, \ldots, z_n$

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \Phi(w_j - w_k) z_j z_k^* \geq 0$$

Linear transformation: if $Y = aX + b$, then

$$\Phi_y(\omega) = e^{\mathrm{i}b\omega}\Phi_x(a\omega)$$

Gaussian variables: let $X \sim \mathcal{N}(\mu, \sigma^2)$

the characteristic function of X is

$$\Phi(\omega) = e^{\mathrm{i}\mu\omega} \cdot e^{-\sigma^2\omega^2/2}$$

(more details of applying CF to show the central limit theorem)

Binomial variables: parameters are n, p and $q = 1 - p$

$$\Phi(\omega) = (pe^{\mathrm{i}\omega} + q)^n$$

Poisson variables: with parameter λ

$$\Phi(\omega) = e^{\lambda(e^{\mathrm{i}\omega} - 1)}$$

Markov and Chebyshev Inequalities

Markov inequality

let X be a *nonnegative* RV with mean $\mathbf{E}[X]$

$$P(X \geq a) \leq \frac{\mathbf{E}[X]}{a}, \quad a > 0$$

Chebyshev inequality

let X be an RV with mean μ and variance σ^2

$$P\left(|X - \mu| \geq a\right) \leq \frac{\sigma^2}{a^2}$$

example: manufacturing of low grade resistors

- assume the averge resistance is 100 ohms (measured by a statistical analysis)
- some of resistors have different values of resistance

if all resistors over 200 ohms will be discarded, what is the maximum fraction of resistors to meet such a criterion ?

using Markov inequality with $\mu = 100$ and $a = 200$

$$P(X \geq 200) \leq \frac{100}{200} = 0.5$$

the percentage of discarded resistors cannot exceed 50% of the total if the variance of the resistance is known to equal 100, find the probability that the resistance values are between 50 and 150

$$\begin{aligned} P(50 \leq X \leq 150) &= P(|X - 100| \leq 50) \\ &= 1 - P(|X - 100| \geq 50) \end{aligned}$$

by Chebyshev inequality

$$P(|X - 100| \geq 50) \;\; \leq \;\; \frac{\sigma^2}{(50)^2} \; = \; 1/25$$

hence,

$$P(50 \leq X \leq 150) \geq 1 - \frac{1}{25} = \frac{24}{25}$$

Chernoff bound

the Chernoff bound is given by

$$P(X \geq a) \quad \leq \quad \inf_{t \geq 0} \mathbf{E}[e^{t(X-a)}]$$

which can be expressed as

$$\log P(X \geq a) \quad \leq \quad \inf_{t \geq 0} \{-ta + \log \mathbf{E} e^{tX}\}$$

- $\mathbf{E}[e^{tX}]$ is the *moment generating function*
- $\log \mathbf{E} e^{tX}$ is called the *cumulant generating function*
- Chernoff bound is useful when $\mathbf{E} e^{tX}$ has an analytical expression

Example: X is Gaussian with zero mean and unit variance

the cumulant generating function is

$$\log \mathbf{E}[e^{tX}] = t^2/2$$

hence,

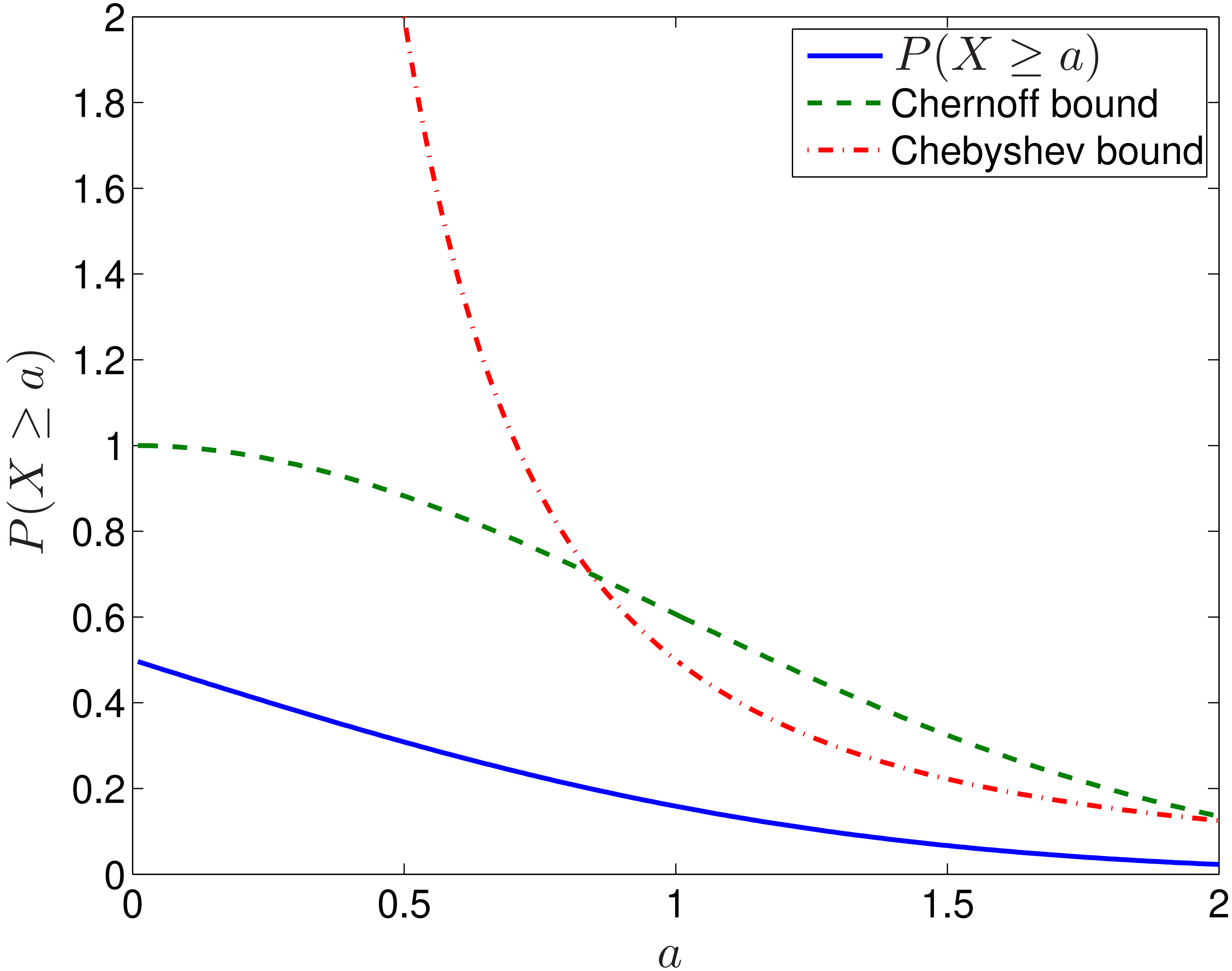
$$\log P(X \geq a) \quad \leq \quad \inf_{t \geq 0} \{-ta + t^2/2\} = -a^2/2$$

and the Chernoff bound gives

$$P(X \geq a) \leq e^{-a^2/2}$$

which is tighter than the Chebyshev inequality:

$$P(|X| \geq a) \leq 1/a^2 \quad \Longrightarrow \quad P(X \geq a) \leq 1/2a^2$$



when a is small, Chebyshev bound is useless while the Chernoff bound is tighter

References

Chapter 3,4 in
A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*, 3rd edition, Pearson Prentice Hall, 2009