- The cdf of normal distribution does not have a close form.
- Theorem. The mean and variance of a $N(\mu, \sigma^2)$ distribution are μ and σ^2, respectively. Why? Check the graphs in LNp.6-31 & ★ in LNp.6-30

Intuition $\quad \sigma^2 = E(X^2)-[E(X)]^2$

  - μ: location parameter; σ (or σ^2): scale (or dispersion) parameter

Proof. $E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \overset{\text{by } (*) \text{ in LNp.6-30}}{=} \int_{-\infty}^{\infty} (\sigma y+\mu)\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy$

$$= \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\, dy}_{-e^{-y^2/2}\big|_{-\infty}^{\infty}} + \mu \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}}_{\text{pdf of } N(0,1)}\, dy$$

$$= \frac{\sigma}{\sqrt{2\pi}}\cdot 0 + \mu\cdot 1 = \mu.$$

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

$$= \sigma^2 \underbrace{\int_{-\infty}^{\infty} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy} + \frac{2\mu\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\, dy$$

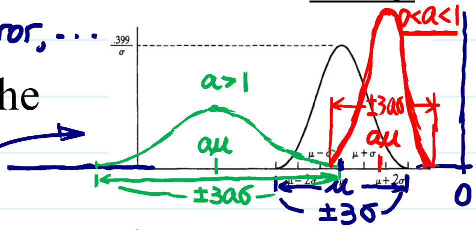
$$+\mu^2 \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}}_{\text{pdf of } N(0,1)}\, dy$$

$$= \sigma^2\cdot 1 + \frac{2\mu\sigma}{\sqrt{2\pi}}\cdot 0 + \mu^2\cdot 1 = \sigma^2+\mu^2.$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{y}_{f} \underbrace{\left(y e^{-\frac{y^2}{2}}\right)}_{g'} dy = \frac{1}{\sqrt{2\pi}} y \left(-e^{-y^2/2}\right)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}}\left(+e^{-\frac{y^2}{2}}\right)}_{\text{pdf of } N(0,1)} dy$$

- Some properties Why? one reason is the central limit thm (CLT)
  - Normal distribution is one of the most widely used distribution. It can be used to model the distribution of many natural phenomena. → e.g. height, weight, error, ...

∵ bell-shaped distribution

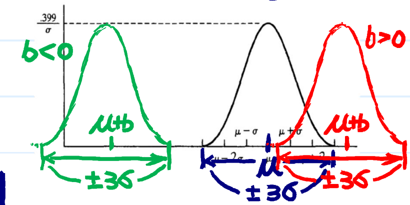


  - Theorem. Suppose that $X \sim N(\mu, \sigma^2)$. The random variable

Recall $E(Y)=aE(X)+b$, $Var(Y)=a^2 Var(X)$

$$Y = aX + b,$$

c.f. graphs in LNp 5-16, 5-18



where $a \neq 0$, is also normally distributed with parameters $a\mu+b$ and $a^2\sigma^2$, i.e.,

$$Y \sim N(a\mu+b,\ a^2\sigma^2).$$

by the example in LNp.6-11

Note: $E(Z)=0$, $Var(Z)=1$ for any r.v. X. But, X & Z may not belong to same distribution in general.

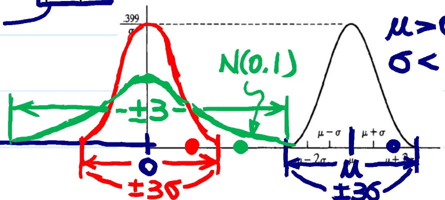
Proof. $f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|} = \frac{1}{\sqrt{2\pi}\underbrace{|a|\sigma}_{\sigma_Y}} e^{-\frac{[y-\overbrace{(a\mu+b)}^{\mu_Y}]^2}{\underbrace{2\sigma^2 a^2}_{\sigma_Y^2}}}.$

  - Corollary. If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X-\mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$$

standardization (check ★ in LNp.6-30) → remove unit

meaning of Z = 1, 1.5, 2?



is a normal random variable with parameters 0 and 1, i.e., $N(0, 1)$, which is called *standard normal distribution*.

- The $N(0, 1)$ distribution is very important since properties of any other normal distributions can be found from those of the standard normal.
  - *The cdf of $N(0, 1)$ is usually denoted by Φ.* — no close form, $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\,dt$
  - Theorem. Suppose that $X \sim N(\mu, \sigma^2)$. The cdf of X is

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

$P(X \le x) = P\left(\underbrace{\frac{X-\mu}{\sigma}}_{Z} \le \frac{x-\mu}{\sigma}\right)$

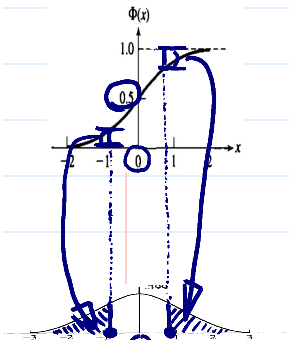
Proof. $F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$ (Z: $N(0,1)$)

  - Example. Suppose that $X \sim N(\mu, \sigma^2)$. For $-\infty < a < b < \infty$,

$$\begin{aligned} P(a < X < b) &= P\left(\frac{a-\mu}{\sigma} < \underbrace{\frac{X-\mu}{\sigma}}_{=Z} < \frac{b-\mu}{\sigma}\right) \\ &= P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) \quad (Z \sim N(0,1)) \\ &= P\left(Z < \frac{b-\mu}{\sigma}\right) - P\left(Z < \frac{a-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

  - Table 5.1 in textbook gives values of Φ. To read the table:

$\Phi(x)$: cdf of $N(0,1)$



z 0 -z

$\phi(x)$: pdf of $N(0,1)$

$(\Phi(z)+\Phi(-z))/2=0.5$

$\Phi(0)=\frac{1}{2}$

$\Phi(0.22)=0.5871$

$\Phi(3.36)=0.9996$

1. Find the first value of x up to the first place of decimal in the left hand column.
2. Find the second place of decimal across the top row.
3. The value of $\Phi(x)$ is where the row from the first step and the column from the second step intersect.

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF x

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
			•	•	•					
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

- For the values greater than z=3.49, $\Phi(z) \approx 1$.
- For negative values of z, use $\Phi(z)=1-\Phi(-z)$

$Z = X_1 + X_2 + \cdots + X_n$, n is large (e.g., binomial ↔ Bernoulli, negative binomial ↔ geometric, gamma ↔ exponential)

$\overset{d}{\approx}$

- Normal distribution plays a central role in the limit theorems of probability (e.g., Central Limit Theorem, CLT, chapter 8) 中央極限定理

an example of CLT

Normal approximation to the Binomial

(LNp.5-43~44, $N_i \uparrow \infty$, $R_i \uparrow \infty$, $\frac{R_i}{N_i} \to p$)

(LNp.5-31~32, $n \uparrow \infty$, $p_n \downarrow 0$, $\lambda = np_n$)

- Recall. Poisson & Hypergeometric approximations to Binomial

standardization

sum of n independent Bernoulli(p)

- Theorem. Suppose that $X_n \sim$binomial(n, p). Define $Z_n = (X_n - np)/\sqrt{np(1-p)}$.

$\mu_n = E(X_n) = np$

$\sigma_n^2 = Var(X_n) = np(1-p)$

$E(Z_n) = 0$

$Var(Z_n) = 1$

$X_n = \sigma_n Z_n + \mu_n$

p is fixed

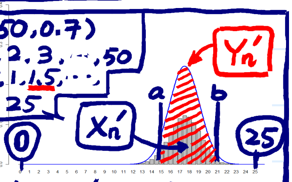
Then, as $n \to \infty$, the distribution of Z_n converge to the $N(0, 1)$ distribution, i.e.,

$$\underbrace{F_{Z_n}(z)}_{\text{cdf of } Z_n} = P(Z_n \le z) \to \underbrace{\Phi(z)}_{\text{cdf of } N(0,1)}.$$

$X_n \sim B(50, 0.7)$

$X_n = 0, 1, 2, 3, \cdots, 50$

$X_n' = 0, 0.5, 1, 1.5, \cdots, 24.5, 25$



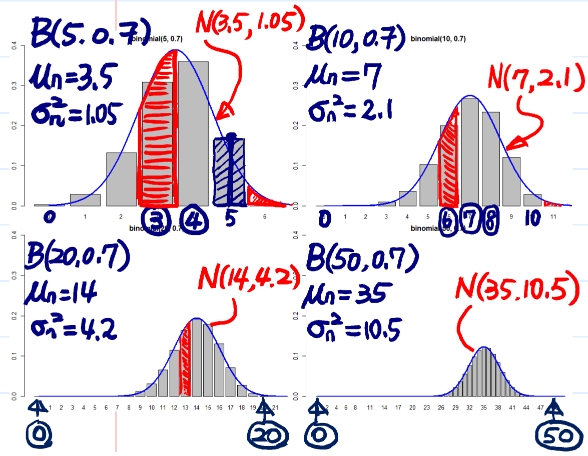
Q: $X_n' = X_n/2$, $Y_n' = Y_n/2$.

pmf $\approx$ pdf ? (No)

cdf $\approx$ cdf ? (Yes)

$P(a < X_n' \le b) = ?$

Proof. It is a special case of the CLT in Chapter 8.



- Plot the pmf of $X_n \sim$binomial(n, p)
- Superimpose the pdf of $Y_n \sim N(\mu_n, \sigma_n^2)$ with $\mu_n = np$ and $\sigma_n^2 = np(1-p)$.
- When n is sufficiently large, the normal pdf approximates the binomial pmf.
- $Z_n \overset{d}{\approx} \underbrace{(Y_n - \mu_n)/\sigma_n}_{\sim N(0,1)}$ (c.f. $Z_n \approx (Y_n - \mu_n)/\sigma_n$)

- The size of n to achieve a good approximation depends on the value of p.

Q: Poisson(λ) $\overset{d}{\approx}$ Normal ?

Ans. Yes, when λ is large. (λ = np)

Why? Check LNp.5-22 graphs.

- For p near 0.5 $\Rightarrow$ moderate n is enough
- For p close to zero or one (skewed) $\Rightarrow$ require much larger n

- Continuity Correction (for integer-valued discrete r.v.'s)
  - **Q**: Why need continuity correction?

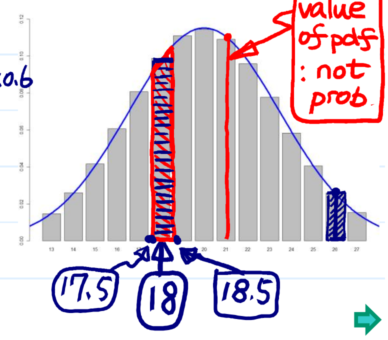
Ans. The binomial(n, p) is a discrete r.v. and we are approximating it with a continuous r.v.. (cf. Normal)

(height) pmf: value $\to$ prob.

pdf: area $\to$ prob.

- For example, suppose $X \sim$binomial(50, 0.4) and we want to find $P(X=18)$, which is larger than 0.
- With the normal r.v. $Y \sim N(20, 12)$, (50×0.4, $50 \times 0.4 \times 0.6$) however, $P(Y=18)=0$ because Y has a continuous distribution

How about using the value of pdf $f_Y(18)$?



- Instead, we make a continuity correction,

$$\begin{aligned}
P(X=18) &= P(17.5 < X < 18.5) \\
&= P\left(\frac{17.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}} < Z_n < \frac{18.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}}\right) \\
&\approx P\left(\frac{17.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}} < \underbrace{Z}_{N(0,1)} < \frac{18.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}}\right) \quad \text{by CLT} \\
&= P\left(-\frac{2.5}{\sqrt{12}} < Z < -\frac{1.5}{\sqrt{12}}\right) = P\left(Z < -\frac{1.5}{\sqrt{12}}\right) - P\left(Z < -\frac{2.5}{\sqrt{12}}\right) \\
&= \Phi\left(-\frac{1.5}{\sqrt{12}}\right) - \Phi\left(-\frac{2.5}{\sqrt{12}}\right) = \left(1-\Phi\left(\frac{1.5}{\sqrt{12}}\right)\right) - \left(1-\Phi\left(\frac{2.5}{\sqrt{12}}\right)\right) \\
&= \Phi\left(2.5/\sqrt{12}\right) - \Phi\left(1.5/\sqrt{12}\right)
\end{aligned}$$

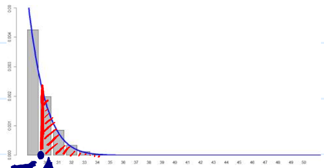
and can obtain the approximate value from Table 5.1.

- Similary,

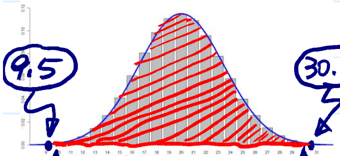
$$P(X \geq 30) = P(X > 29.5) = P\left(Z_n > \frac{29.5-(50\cdot 0.4)}{\sqrt{12}}\right)$$

$$\approx P(\underbrace{Z}_{N(0,1)} > 9.5/\sqrt{12}) = 1 - \Phi(9.5/\sqrt{12}). \quad \text{by CLT}$$

and



$$\begin{aligned}
P(10 \leq X \leq 30) &= P(9.5 < X < 30.5) \\
&= P\left(\frac{9.5-(50\cdot 0.4)}{\sqrt{12}} < Z_n < \frac{30.5-(50\cdot 0.4)}{\sqrt{12}}\right) \\
&\approx P\left(-10.5/\sqrt{12} < \underbrace{Z}_{N(0,1)} < 10.5/\sqrt{12}\right) \quad \text{by CLT} \\
&= \Phi\left(10.5/\sqrt{12}\right) - \underbrace{\Phi\left(-10.5/\sqrt{12}\right)}_{=1-\Phi\left(\frac{10.5}{\sqrt{12}}\right)} \\
&= 2\cdot\Phi\left(10.5/\sqrt{12}\right) - 1
\end{aligned}$$



➢ Summary for $X \sim$ Normal(μ, σ^2)

- Pdf: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$
- Cdf: no close form, but usually denoted by $\Phi((x-\mu)/\sigma)$.
- Parameters: $\mu \in \mathbb{R}$ and $\sigma > 0$.
- Mean: $E(X) = \mu$.
- Variance: $Var(X) = \sigma^2$.

章值

• Weibull Distribution

➢ For $\alpha, \beta > 0$ and $\nu \in \mathbb{R}$, the function

fixed constants

$$f(x) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^{\beta}}, & \text{if } x \geq \nu, \\ 0, & \text{if } x < \nu, \end{cases}$$

possible values of X

is a pdf since (1) $f(x) \geq 0$ for all $x \in \mathbb{R}$, and (2)

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{\nu}^{\infty} \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^{\beta}}\,dx$$

$$\stackrel{(\Delta)}{=} \int_0^{\infty} \underbrace{e^{-y}}_{\text{pdf of exponential(1)}}\,dy = -e^{-y}\Big|_0^{\infty} = 1.$$

$$y = \left(\frac{x-\nu}{\alpha}\right)^{\beta} \Rightarrow x = \alpha y^{\frac{1}{\beta}} + \nu$$

$$\frac{dx}{dy} = \frac{\alpha}{\beta} y^{\frac{1}{\beta}-1} \Rightarrow dx = \frac{\alpha}{\beta} y^{\frac{1}{\beta}-1}\,dy$$

(Δ)

- The distribution of a random variable X with this pdf is called the *Weibull* distribution with parameters α, β, and ν.

➢(exercise) The cdf of Weibull distribution is

by (Δ) in LNp.6-39

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta}, & \text{if } x \geq \nu, \\ 0, & \text{if } x < \nu. \end{cases}$$

➢Theorem. The mean and variance of a Weibull distribution with parameters α, β, and ν are

$$\mu = \alpha\Gamma\left(1 + \tfrac{1}{\beta}\right) + \nu \quad \text{and}$$

$$E(X^2)-[E(X)]^2 = \sigma^2 = \alpha^2 \left\{\Gamma\left(1 + \tfrac{2}{\beta}\right) - \left[\Gamma\left(1 + \tfrac{1}{\beta}\right)\right]^2\right\}.$$

Proof. $E(X) = \int_v^\infty x \cdot \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta}\, dx$

$$= \int_0^\infty (\alpha y^{1/\beta} + \nu) e^{-y}\, dy$$

$$= \alpha \int_0^\infty y^{(1/\beta+1)-1} e^{-y}\, dy + \nu \int_0^\infty \underbrace{e^{-y}}_{\text{pdf of exponential(1)}}\, dy = \alpha\Gamma\left(\tfrac{1}{\beta} + 1\right) + \nu$$

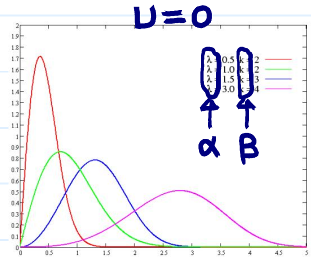
$$E(X^2) = \int_v^\infty x^2 \cdot \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta}\, dx$$

$$= \int_0^\infty (\alpha y^{1/\beta} + \nu)^2 e^{-y}\, dy$$

$$= \alpha^2 \int_0^\infty y^{(2/\beta+1)-1} e^{-y}\, dy + 2\alpha\nu \int_0^\infty y^{(1/\beta+1)-1} e^{-y}\, dy + \nu^2 \int_0^\infty e^{-y}\, dy$$

$$= \alpha^2\Gamma\left(\tfrac{2}{\beta} + 1\right) + 2\alpha\nu\Gamma\left(\tfrac{1}{\beta} + 1\right) + \nu^2$$

➢Some properties



- Weibull distribution is widely used to model lifetime (cf., exponential) memoryless
- α: scale parameter; β: shape parameter; ν: location parameter
- Theorem. If X~exponential(λ), then

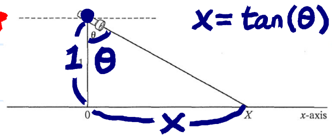
Hint: $X \sim$ Gamma(α, λ) ⇒ $aX \sim$ Gamma(α, $\frac{\lambda}{a}$), $a>0$ (LNp.6-26)

$$Y = \alpha\,(\lambda X)^{1/\beta} + \nu$$

Note: $\lambda X \sim$ exponential(1)

is distributed as Weibull with parameters α, β, and ν (exercise). Thm(LNp.6-10)

• Cauchy Distribution 柯西



➢For μ∈ℝ and σ>0, the function

fixed constants

possible values of X

$$f(x) = \frac{\sigma}{\pi} \cdot \frac{1}{\sigma^2 + (x-\mu)^2}, \quad -\infty < x < \infty,$$

θ~Uniform($-\frac{\pi}{2}, \frac{\pi}{2}$)
Let X=tan(θ), then
X~Cauchy(0,1)
(exercise)

(∇)
$y = \frac{x-\mu}{\sigma}$
⇒ $x = \sigma y + \mu$
$\frac{dx}{dy} = \sigma$
⇒ $dx = \sigma dy$

is a pdf since (1) $f(x) \geq 0$ for all x∈ℝ, and (2)

$$\int_{-\infty}^\infty f(x)\, dx = \int_{-\infty}^\infty \frac{\sigma}{\pi} \frac{1}{\sigma^2 + (x-\mu)^2}\, dx$$

$$= \int_{-\infty}^\infty \underbrace{\frac{1}{\pi}\frac{1}{1+y^2}}_{\text{pdf of Cauchy(0,1)}}\, dy = \frac{1}{\pi}\tan^{-1}(y)\Big|_{-\infty}^{\infty} = 1.$$

- The distribution of a random variable X with this pdf is called the *Cauchy* distribution with parameters μ and σ, denoted by Cauchy(μ, σ).

  not mean — not standard deviation

Recall. The 2 Thms in LNp.6-8 & 6-9, and ★ in LNp.6-41.

➢ The cdf of Cauchy distribution is

$$F(x)=\int_{-\infty}^{x}\frac{\sigma}{\pi}\frac{1}{\sigma^2+(y-\mu)^2}\,dy=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}\left(\frac{x-\mu}{\sigma}\right)$$

for $-\infty<x<\infty$. (exercise) by (▽) in LNp.6-41.

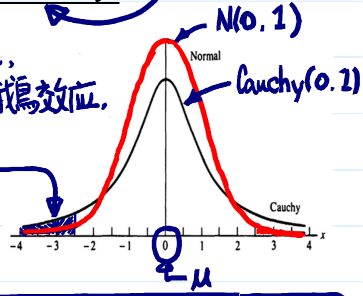
➢ The mean and variance of Cauchy distribution do *not exist* because the integral does *not converge absolutely*

$\because \int_{-\infty}^{\infty}|x|f(x)dx=\infty$, $\int_{-\infty}^{\infty}x^2 f(x)dx=\infty$

➢ Some properties

Why? check the graph ★ in LNp.6-41

- Cauchy is a heavy tail distribution — model traffic, wealth, 黑天鵝效應, ...



- μ: location parameter; σ: scale parameter

c.f. Normal distribution

- Theorem. If X~Cauchy(μ, σ), then $aX+b$~Cauchy($a\mu+b$, $|a|\sigma$).

  Proof. (exercise) Thm (LNp.6-10)

Note: a pdf $f(x)\downarrow 0$, when $x\to\infty$ or $x\to-\infty$. But, $f(x)\downarrow 0$ how fast?

❖ **Reading**: textbook, Sec 5.4, 5.5, 5.6