

Continuous

① Gamma

$$\Gamma(\alpha, \lambda) \Rightarrow f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

Mean

$$\begin{aligned} \mathbb{E}[X] &= \int \frac{\lambda^\alpha}{\Gamma(\alpha)} x^\alpha e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} \end{aligned}$$

$$= \frac{\alpha}{\lambda}$$

2-nd moment

$$\mathbb{E}[X^2] = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}}$$

$$= \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\Rightarrow \text{Var}[X] = \frac{\alpha}{\lambda^2}$$

$$\Gamma(\frac{1}{2}) = \int_0^\infty x^{-\frac{1}{2}} e^{-x} dx$$

$$= \int_0^\infty \frac{1}{\sqrt{x}} e^{-x} dx$$

$$x=u^2 \Rightarrow \int_0^\infty e^{-u^2} du \rightarrow N(0, \frac{1}{2})$$

$$= \sqrt{2\pi} \cdot \frac{\sqrt{2}}{2}$$

$$= \sqrt{\pi}$$

a) Gamma(1, λ) $\Rightarrow \exp(\lambda)$

b) Gamma($\frac{n}{2}, \frac{1}{2}$) $\Rightarrow \chi^2(n)$

② Beta distribution

$$\text{Beta}(a, b) \rightarrow f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \quad \mathcal{X} = [0, 1]$$

$$\mathbb{E}[X] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}$$

$$= \frac{a}{a+b}$$

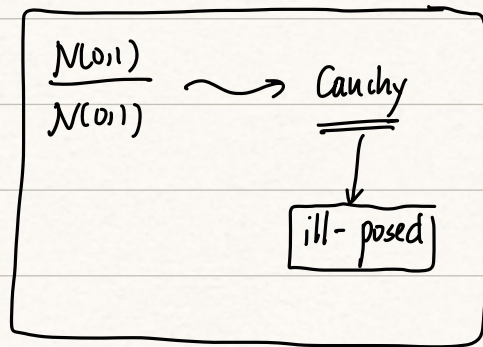
$$\mathbb{E}[X^2] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)}$$

$$= \frac{a \cdot (a+1)}{(a+b)(a+b+1)}$$

a) Beta(1,1) = Uniform(0,1)

b) $\begin{cases} a \longleftrightarrow 1 \\ b \longleftrightarrow 1 \end{cases} \Rightarrow \underline{\text{不同单调性 } f(x)}$

③ Cauchy dist



!! critical problem

④ Double exponential distribution

$$f(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$$

$$E[X] = \mu$$

$$\text{Var}[X] = 2b^2$$

Location-Scale family

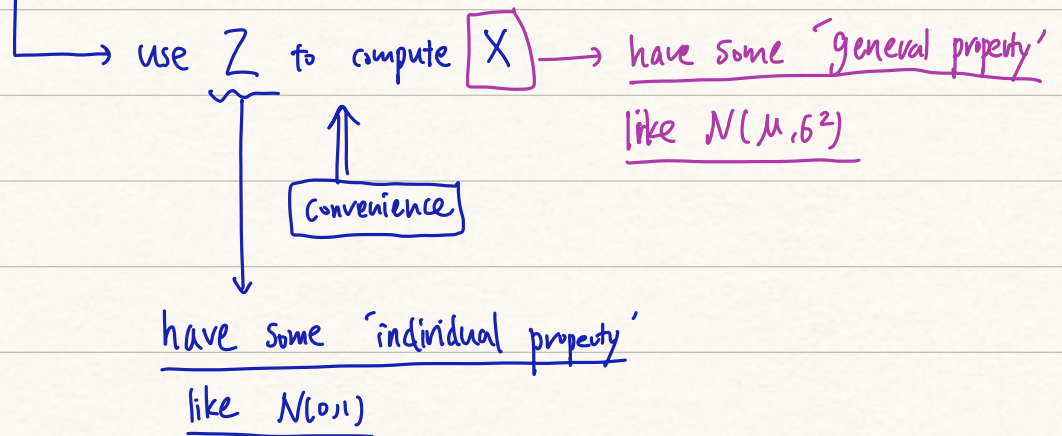
$$\begin{cases} \mu \rightarrow \text{location parameter} \\ b \rightarrow \text{scale parameter} \end{cases}$$

① if $f(x)$ is a p.d.f.

then $\frac{1}{b} f(\frac{x-\mu}{b})$ is also a valid pdf

② $Z \sim f(x), X \sim \frac{1}{b} f(\frac{x-\mu}{b})$

$\Leftrightarrow \underline{X = bZ + \mu}$



11 | 21 More details about Beta(a, b) distribution!

① Expression:

$$X \sim \text{Beta}(a, b)$$

$$\Rightarrow f_X(x) = \frac{1}{B(a, b)} \int_0^1 x^{a-1} (1-x)^{b-1} dx$$
$$= \frac{\Gamma(a+b)}{\Gamma(a) \Gamma(b)} \int_0^1 x^{a-1} (1-x)^{b-1} dx$$

Here, $B(a, b)$ is the normalization term

$$\Leftrightarrow \underline{B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx}$$

② Interpretation

a) check the expectation & variance

$$E[X] = \frac{\Gamma(a+b)}{\Gamma(a) \Gamma(b)} \int_0^1 x^a (1-x)^{b-1} dx$$

$$= \frac{\Gamma(a+b)}{\Gamma(a) \Gamma(b)} \frac{\Gamma(a+1) \Gamma(b)}{\Gamma(a+b+1)}$$

$$= \frac{a}{a+b}$$

$$\mathbb{E}[X^2] = \frac{a(a+1)}{(a+b)(a+b+1)} \quad \frac{a^2(a+b) - a^2(a+b+1) + a^2 + ab}{(a+b)(a+b+1)}$$

$$\begin{aligned} \text{Var}[X] &= \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a \cdot a}{(a+b)(a+b)} \\ &= \frac{a(a+1)(a+b) - a^2(a+b+1)}{(a+b)^2(a+b+1)} = \frac{ab}{(a+b)^2(a+b+1)} \end{aligned}$$

To conclude: $\begin{cases} \mathbb{E}[X] = \frac{a}{a+b} \\ \text{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)} \end{cases}$ and Support Set of X is [0,1]

★

b) $X \rightarrow$ view as the random probability

whose $\begin{cases} \frac{a}{a+b} \text{ can be modeled as Success Probability} \\ \text{if } a, b \text{ becomes bigger, then uncertainty shrinks!} \end{cases}$

(Variance becomes smaller)

[e.g.] Beta(2,2) versus Beta(4,4)

↓
more skewed \leftrightarrow large variance

③ Advantage: Conjugate Prior w.r.t many likelihood

↓
Binomial / NegBin

[e.g.] $\begin{cases} X | \lambda = \lambda \sim \text{Binomial}(n, \lambda) \\ \lambda \sim \text{Beta}(a, b) \end{cases}$

$$\begin{aligned} P_{\lambda|X}(x|k) &\propto P_{X|\lambda}(k|\lambda) \cdot P_{\lambda}(\lambda) \\ &\propto \binom{n}{k} \lambda^k (1-\lambda)^{n-k} \cdot \lambda^{a-1} (1-\lambda)^{b-1} \end{aligned}$$

$$\propto \lambda^{(a+k-1)} (1-\lambda)^{(n-k+b-1)}$$

$$\Rightarrow \Lambda | X=k \sim \text{Beta}(a+k, b+n-k)$$

④ Order-Statistic \longleftrightarrow Beta Dist

a) $X_1, \dots, X_n \sim F$

\rightarrow what is the distribution of $X_{(k)}$?

$$\begin{aligned} & P(X_{(k)} \in [x, x+\Delta x]) \\ &= \binom{n}{k-1, k, n-k-1} \cdot (F(x))^k \cdot (F(x+\Delta x) - F(x)) \cdot (1 - F(x+\Delta x))^{n-k-1} \end{aligned}$$

$$= \frac{n!}{(k-1)! \cdot k! \cdot (n-k-1)!} F(x)^k (F(x+\Delta x) - F(x)) (1 - F(x+\Delta x))^{n-k-1}$$

$$= p_{X_{(k)}}(x) \Delta x + o(\Delta x)$$

$$\Rightarrow \underline{p_{X_{(k)}}(x) = \frac{n!}{k! (n-k-1)!} F(x)^k f(x) (1 - F(x))^{n-k-1}}$$

b) when $X_1, \dots, X_n \sim \text{Uniform}[0,1]$

$$\text{then } p_{U_{(k)}}(x) = \frac{n!}{k! (n-k-1)!} x^{k+1-1} (1-x)^{n-k-1}$$

$$= \frac{\Gamma(n+1)}{\Gamma(k+1) \Gamma(n-k)} x^{k+1-1} (1-x)^{n-k-1}$$

that is $U_{(k)} \sim \text{Beta}(k+1, n-k)$

\rightarrow intuition about the Random Variable Beta

⑤ Lastly, determine $B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$
normalization constant

Technique \rightarrow integration by parts

Lemma 1

$$\begin{aligned} \rightarrow B(a, b) &= \int_0^1 x^{a-1} (1-x)^{b-1} dx \\ &= - \int_0^1 x^{a-1} d\left(\frac{1}{b} (1-x)^b\right) \end{aligned}$$

$$= - x^{a-1} \frac{1}{b} (1-x)^b \Big|_0^1 + \int_0^1 \frac{a-1}{b} x^{a-2} (1-x)^b dx$$

$$= \frac{a-1}{b} B(a-1, b+1)$$

Lemma 2

$$\begin{aligned} B(1, k) &= \int_0^1 (1-x)^{k-1} dx \\ &= - \frac{1}{k} (1-x)^k \Big|_0^1 = \frac{1}{k} \end{aligned}$$

$$\text{Therefore, } B(a, b) = \frac{a-1}{b} B(a-1, b+1)$$

$$= \frac{a-1}{b} \cdots \frac{1}{a+b-2} B(1, a+b-1)$$

$$= \frac{(a-1) \cdots 1}{(a+b-1) \cdots b} = \frac{\Gamma(a)}{\Gamma(a+b)} = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)}$$