Lecture 11: Useful distributions (part II)

Gamma distribution *gamma*(α,β)

Let $\alpha > 0$ and $\beta > 0$ be two constants.

The gamma distribution with shape parameter α and scale parameter β has pdf

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & x \le 0 \end{cases}$$

where

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

is called the gamma function,

$$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(\alpha+1) = \alpha\Gamma(\alpha), \quad \Gamma(n) = (n-1)!,\ n = 1,2,...$$

If $\alpha < 1$, then the pdf of *gamma*(α,β) is decreasing in x and unbounded at $x = 0$.

If $\alpha \ge 1$, then the pdf of *gamma*(α,β) is bounded, increasing in $x < (\alpha-1)\beta$, and decreasing in $x > (\alpha-1)\beta$.

Previously, we showed that, if $X \sim$ *gamma*(α,β), then

$$M_X(t) = (1-\beta t)^{-\alpha}, \quad t < 1/\beta$$
$$E(X^n) = \alpha(\alpha+1)\cdots(\alpha+n-1)\beta^n, \quad n = 1, 2, ...$$

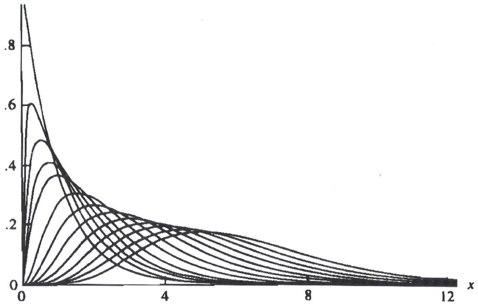
In particular,

$$E(X) = \alpha\beta, \quad E(X^2) = \alpha(\alpha+1)\beta^2, \quad \mathrm{Var}(X) = \alpha\beta^2$$

From the mfg, the chf is

$$\phi_X(t) = (1-i\beta t)^{-\alpha}, \quad t \in \mathscr{R}$$

Some gamma pdf curves



Chi-square distribution

The gamma distribution *gamma*(α, β) with $\beta = 2$ and $\alpha = k/2$ for a positive integer k is called the chi-square distribution with k degrees of freedom, and its pdf is

$$f(x) = \begin{cases} \frac{1}{\Gamma(k/2)2^{k/2}} x^{(k/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

- If $X \sim f$, then
$$E(X) = k, \quad \mathrm{Var}(X) = 2k,$$
and the mgf $M_X(t)$ and chf $\phi_X(t)$ are given by
$$M_X(t) = (1-2t)^{-k/2}, \;\; t < 1/2, \quad \phi_X(t) = (1-2it)^{-k/2}, \;\; t \in \mathscr{R}$$
- The chi-square distribution is closely related to the normal distribution.
- In Example 2.1.9, we showed that if $X \sim N(0,1)$, then $X^2 \sim$ chi-square with 1 degree of freedom.
- A similar result about the chi-square distribution with k degrees of freedom will be introduced later.

Exponential distribution *exponential*(β)

A special case of the gamma distribution with $\alpha = 1$ is called the exponential distribution, i.e., $gamma(1,\beta) = exponential(\beta)$.

The $exponential(\beta)$ distribution has pdf

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

For any $x > 0$,

$$P(X \leq x) = \frac{1}{\beta} \int_0^x e^{-t/\beta} dt = e^{-t/\beta} \Big|_x^0 = 1 - e^{-x/\beta}$$

Hence, the cdf of $exponential(\beta)$ has an explicit form

$$F(x) = \begin{cases} 1 - e^{-x/\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

An example of $X \sim exponential(\beta)$ is the lifetime.

Since *exponential*$(\beta) =$ *gamma*$(1,\beta)$, if $X \sim$ *exponential*(β),

$$E(X) = \beta, \quad \mathrm{Var}(X) = \beta^2, \quad M_X(t) = \frac{1}{1-\beta t}, \quad t < \frac{1}{\beta},$$

$$\phi_X(t) = \frac{1}{1-i\beta t}, \quad t \in \mathscr{R}$$

The exponential distribution has a memoryless property, i.e.,

$$P(X > s | X > t) = P(X > s-t)$$

for any $s > t \geq 0$, because

$$\begin{aligned} P(X > s | X > t) &= \frac{P(X > s, X > t)}{P(X > t)} = \frac{P(X > s)}{P(X > t)} \\ &= \frac{1-F(s)}{1-F(t)} = \frac{e^{-s/\beta}}{e^{-t/\beta}} = e^{-(s-t)/\beta} \\ &= 1 - F(s-t) = P(X > s-t) \end{aligned}$$

That means, if X is the lifetime of a product, then the lifetime of a product that has been used t hours has the same distribution as that of a new product (new = used).

Two parameter exponential distribution *exponential*(μ,β)

If $X \sim$ *exponential*(β) and $\mu \in \mathscr{R}$ is a constant, then the distribution of $Y = X + \mu$ is called the two parameter exponential distribution and denoted by *exponential*(μ,β).
Its pdf and cdf are (by transformation)

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-(x-\mu)/\beta} & x \geq \mu \\ 0 & x < \mu \end{cases} \qquad F(x) = \begin{cases} 1 - e^{-(x-\mu)/\beta} & x \geq \mu \\ 0 & x < \mu \end{cases}$$

and, if $Y \sim$ *exponential*(μ,β),

$$E(Y)=\mu+\beta,\ \mathrm{Var}(Y)=\beta^2,\ M_Y(t)=\frac{e^{\mu t}}{1-\beta t},\ t<\frac{1}{\beta},\ \phi_Y(t)=\frac{e^{i\mu t}}{1-i\beta t},\ t\in\mathscr{R}$$

Double exponential distribution *double-exponential*(μ,σ)

By reflecting the pdf of *exponential*(μ,σ) around μ, we obtain the *double-exponential*(μ,σ) pdf that is symmetric about μ:

$$f(x) = \begin{cases} \frac{1}{2\sigma} e^{-(x-\mu)/\sigma} & x \geq \mu \\ \frac{1}{2\sigma} e^{(x-\mu)/\sigma} & x < \mu \end{cases} = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \quad x \in \mathscr{R}$$

Two parameter exponential distribution *exponential*(μ,β)

If $X \sim$ *exponential*(β) and $\mu \in \mathscr{R}$ is a constant, then the distribution of $Y = X + \mu$ is called the two parameter exponential distribution and denoted by *exponential*(μ,β).
Its pdf and cdf are (by transformation)

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-(x-\mu)/\beta} & x \geq \mu \\ 0 & x < \mu \end{cases} \qquad F(x) = \begin{cases} 1 - e^{-(x-\mu)/\beta} & x \geq \mu \\ 0 & x < \mu \end{cases}$$

and, if $Y \sim$ *exponential*(μ,β),

$$E(Y) = \mu + \beta, \ \mathrm{Var}(Y) = \beta^2, \ M_Y(t) = \frac{e^{\mu t}}{1-\beta t}, \ t < \frac{1}{\beta}, \ \phi_Y(t) = \frac{e^{i\mu t}}{1-i\beta t}, \ t \in \mathscr{R}$$

Double exponential distribution *double-exponential*(μ,σ)

By reflecting the pdf of *exponential*(μ,σ) around μ, we obtain the *double-exponential*(μ,σ) pdf that is symmetric about μ:

$$f(x) = \begin{cases} \frac{1}{2\sigma} e^{-(x-\mu)/\sigma} & x \geq \mu \\ \frac{1}{2\sigma} e^{(x-\mu)/\sigma} & x < \mu \end{cases} = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \quad x \in \mathscr{R}$$

- This pdf is not bell-shaped; in fact, it has a peak (a non-differentiable point) at $x = \mu$.
- Its cdf is

$$F(x) = \begin{cases} 1 - \frac{1}{2} e^{-(x-\mu)/\sigma} & x \geq \mu \\ \frac{1}{2} e^{(x-\mu)/\sigma} & x < \mu \end{cases}$$

- If $X \sim$ *double-exponential*(μ, σ), then $Z = (X - \mu)/\sigma \sim$ *double-exponential*$(0, 1)$.
- If $Z = (X - \mu)/\sigma \sim$ *double-exponential*$(0, 1)$, then

$$E(Z) = \frac{1}{2} \int_{-\infty}^{\infty} x e^{-|x|} dx = 0$$

because $xe^{-|x|}$ is an odd function, and

$$\mathrm{Var}(Z) = E(Z^2) = \frac{1}{2} \int_{-\infty}^{\infty} x^2 e^{-|x|} dx = \int_0^{\infty} x^2 e^{-x} dx = \Gamma(3) = 2$$

- If $X \sim$ *double-exponential*(μ, σ), then $X = \sigma Z + \mu$, $Z = (X - \mu)/\sigma \sim$ *double-exponential*$(0, 1)$, and

$$E(X) = E(\sigma Z + \mu) = \mu, \qquad \mathrm{Var}(X) = \mathrm{Var}(\sigma Z + \mu) = \sigma^2 \mathrm{Var}(Z) = 2\sigma^2$$

Cauchy distribution $Cauchy(\mu, \sigma)$

For constants $\mu \in \mathscr{R}$ and $\sigma > 0$, the $Cauchy(\mu,\sigma)$ distribution has pdf

$$f(x) = \frac{\sigma}{\pi[\sigma^2 + (x-\mu)^2]}, \quad x \in \mathscr{R}$$

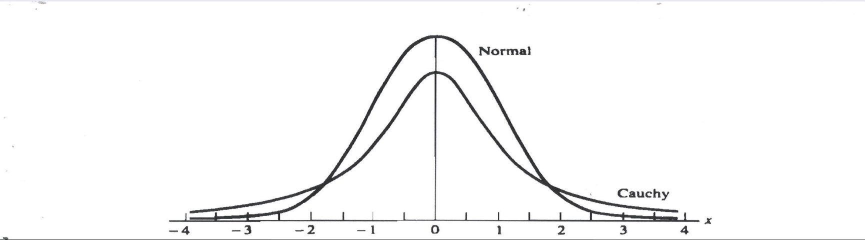
The Cauchy pdf is bell-shaped and symmetric about μ.

Its biggest difference from the normal pdf is that, if $X \sim Cauchy(\mu,\sigma)$, then $E|X| = \infty$ (and hence $E|X|^n = \infty$, $n = 2, 3, ...$)

The mgf of X is ∞ except at 0 and the chf of X is $e^{i\mu+\sigma|t|}$, $t \in \mathscr{R}$.

μ still measures the center of $Cauchy(\mu,\sigma)$ although the expectation does not exist: μ is the median in the sense that $P(X < \mu) = \frac{1}{2}$.

Standard normal and Cauchy pdf curves



Logistic distribution *logistic*(μ, σ)

For constants $\mu \in \mathscr{R}$ and $\sigma > 0$, the *logistic*(μ, σ) distribution has pdf

$$f(x) = \frac{e^{-(x-\mu)/\sigma}}{\sigma[1 + e^{-(x-\mu)/\sigma}]^2}, \quad x \in \mathscr{R}$$

- This pdf is again bell-shaped and symmetric about μ.
- The cdf of *logistic*(μ, σ) has a close form:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{1 + e^{-(x-\mu)/\sigma}}, \quad x \in \mathscr{R}$$

- By symmetry, $E(X) = \mu$ if $X \sim$ *logistic*(μ, σ).
- The variance of $X \sim$ *logistic*(μ, σ) is not easy to obtain, but we give the result here: $\mathrm{Var}(X) = \sigma^2\pi^2/3$.

Pareto distribution *pareto*(α, β)

For constants $\alpha > 0$ and $\beta > 0$, the *pareto*(α, σ) distribution has pdf

$$f(x) = \begin{cases} \alpha\beta^{\alpha}x^{-(\alpha+1)} & x > \beta \\ 0 & x \le \beta \end{cases}$$

Logistic distribution *logistic*(μ,σ)

For constants $\mu \in \mathscr{R}$ and $\sigma > 0$, the *logistic*(μ,σ) distribution has pdf

$$f(x) = \frac{e^{-(x-\mu)/\sigma}}{\sigma[1+e^{-(x-\mu)/\sigma}]^2}, \quad x \in \mathscr{R}$$

- This pdf is again bell-shaped and symmetric about μ.
- The cdf of *logistic*(μ,σ) has a close form:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{1+e^{-(x-\mu)/\sigma}}, \quad x \in \mathscr{R}$$

- By symmetry, $E(X) = \mu$ if $X \sim$ *logistic*(μ,σ).
- The variance of $X \sim$ *logistic*(μ,σ) is not easy to obtain, but we give the result here: $\mathrm{Var}(X) = \sigma^2\pi^2/3$.

Pareto distribution *pareto*(α,β)

For constants $\alpha > 0$ and $\beta > 0$, the *pareto*(α,σ) distribution has pdf

$$f(x) = \begin{cases} \alpha\beta^{\alpha}x^{-(\alpha+1)} & x > \beta \\ 0 & x \leq \beta \end{cases}$$

- First, f is indeed a pdf, because

$$\int_{-\infty}^{\infty} f(x)dx = \alpha\beta^{\alpha}\int_{\beta}^{\infty} x^{-(\alpha+1)}dx = \beta^{\alpha}x^{-\alpha}\Big|_{\infty}^{\beta} = \beta^{\alpha}\beta^{-\alpha} = 1$$

- Using a similar argument, we can obtain the cdf of $pareto(\alpha, \beta)$ as

$$F(x) = \begin{cases} 1 - \left(\frac{\beta}{x}\right)^{\alpha} & x > \beta \\ 0 & x \le \beta \end{cases}$$

- Since the integral $\int_{\beta}^{\infty} x^{-t}dx$ is finite iff $t > 1$, $E(X) = \infty$ if $\alpha \le 1$ when $X \sim pareto(\alpha, \beta)$; if $\alpha > 1$, then

$$E(X) = \alpha\beta^{\alpha}\int_{\beta}^{\infty} x^{-\alpha}dx = \frac{\alpha\beta^{\alpha}}{\alpha - 1}x^{-(\alpha-1)}\Big|_{\infty}^{\beta} = \frac{\alpha\beta^{\alpha}}{\alpha - 1}\beta^{-(\alpha-1)} = \frac{\alpha\beta}{\alpha - 1}$$

- Similarly, $\mathrm{Var}(X) = \infty$ if $\alpha \le 2$; and if $\alpha > 2$,

$$E(X^2) = \alpha\beta^{\alpha}\int_{\beta}^{\infty} x^{-\alpha+1}dx = \frac{\alpha\beta^{\alpha}}{\alpha - 2}x^{-\alpha+2}\Big|_{\infty}^{\beta} = \frac{\alpha\beta^{\alpha}}{\alpha - 2}\beta^{-\alpha+2} = \frac{\alpha\beta^2}{\alpha - 2}$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{\alpha\beta^2}{\alpha - 2} - \frac{\alpha^2\beta^2}{(\alpha - 1)^2} = \frac{\alpha\beta^2}{(\alpha - 1)^2(\alpha - 2)}$$

Weibull distribution $\textit{Weibull}(\gamma,\beta)$

For constants $\gamma > 0$ and $\beta > 0$, if $X \sim \textit{exponential}(\beta)$, then $Y = X^{1/\gamma} \sim \textit{Weibull}(\gamma,\beta)$ with pdf

$$f(x) = \begin{cases} \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

- An example of $Y \sim \textit{Weibull}(\gamma,\beta)$ is lifetime or failure time.
- If $Y \sim \textit{Weibull}(\gamma,\beta)$, then $X = Y^\gamma \sim \textit{exponential}(\beta)$ and

$$\begin{aligned} E(Y) &= E(X^{1/\gamma}) = \frac{1}{\beta}\int_0^\infty x^{1/\gamma} e^{-x/\beta} dx \\ &= \beta^{1/\gamma} \int_0^\infty u^{1/\gamma} e^u du = \beta^{1/\gamma}\Gamma\left(\frac{1}{\gamma}+1\right) \end{aligned}$$

Similarly, we can obtain that

$$\mathrm{Var}(Y) = \beta^{2/\gamma}\left\{\Gamma\left(\frac{2}{\gamma}+1\right) - \left[\Gamma\left(\frac{1}{\gamma}+1\right)\right]^2\right\}$$

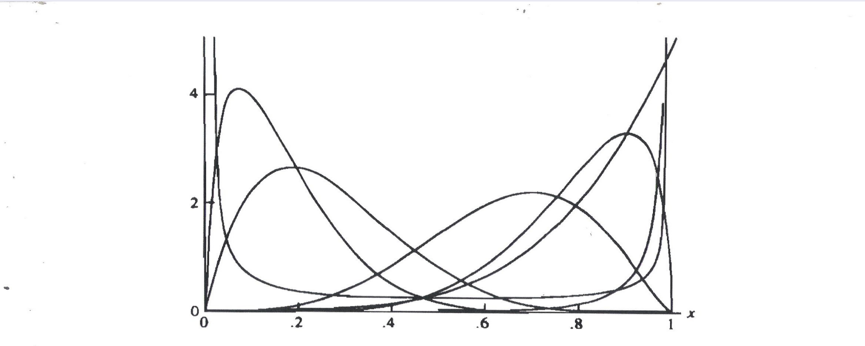
Beta distribution *beta*(α, β)

For constants $\alpha > 0$ and $\beta > 0$, the *beta*(α, β) distribution has pdf

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

This is a pdf because

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$



- Since $\Gamma(2) = \Gamma(1) = 1$, *beta*$(1,1)$ is the same as *uniform*$(0,1)$.
- If $X \sim$ *beta*(α,β), then $1 - X \sim$ *beta*(β,α).
- The pdf of *beta*(α,β) can be increasing $(\alpha > 1, \beta = 1)$, decreasing $(\alpha = 1, \beta > 1)$, U-shaped $(\alpha < 1, \beta < 1)$, or unimodal $(\alpha > 1, \beta > 1)$.
- If $\alpha = \beta$, then the pdf of *beta*(α,β) is symmetric about $\frac{1}{2}$.
- For any $r > 0$, if $X \sim$ *beta*(α,β), then

$$E(X^r) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{r+\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha+\beta)\Gamma(r+\alpha)}{\Gamma(\alpha)\Gamma(r+\alpha+\beta)}$$

In particular,

$$E(X) = \frac{\alpha}{\alpha+\beta}, \qquad E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

Then

$$\mathrm{Var}(X) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Some other continuous distributions will be introduced in Chapter 5.

- Since $\Gamma(2) = \Gamma(1) = 1$, *beta*$(1,1)$ is the same as *uniform*$(0,1)$.
- If $X \sim$ *beta*(α,β), then $1-X \sim$ *beta*(β,α).
- The pdf of *beta*(α,β) can be increasing ($\alpha > 1, \beta = 1$), decreasing ($\alpha = 1, \beta > 1$), U-shaped ($\alpha < 1, \beta < 1$), or unimodal ($\alpha > 1, \beta > 1$).
- If $\alpha = \beta$, then the pdf of *beta*(α,β) is symmetric about $\frac{1}{2}$.
- For any $r > 0$, if $X \sim$ *beta*(α,β), then

$$E(X^r) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{r+\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha+\beta)\Gamma(r+\alpha)}{\Gamma(\alpha)\Gamma(r+\alpha+\beta)}$$

In particular,

$$E(X) = \frac{\alpha}{\alpha+\beta}, \qquad E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

Then

$$\mathrm{Var}(X) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Some other continuous distributions will be introduced in Chapter 5.