目录

序

第 1 章 Special Distributions

1.1 正态分布

定义 1.1. Normal Distribution

A random variable X has the normal distribution with mean and variance 2 (< < and > 0) if X has a continuous distribution with the following $p.d.f.$

$$f(x|\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1.1}$$

for $-\infty<x<\infty$

♣

定理 1.1

正态分布的概率密度函数积分为 1。

$$\int_{-\infty}^{\infty}f(x|\mu,\sigma^2)dx=1 \tag{1.2}$$

♡

证明 let $y=\frac{x-\mu}{\sigma}$, then

$$\int_{-\infty}^{\infty}f(x|\mu,\sigma^2)dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy \tag{1.3}$$

let $I=\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy$, then

$$\begin{aligned}
I^2&=\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz\\
&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{y^2+z^2}{2}}dydz && \text{let } y=r\cos\theta,\ z=r\sin\theta\\
&=\int_0^{2\pi}\int_0^{\infty}e^{-\frac{r^2}{2}}rdrd\theta && \text{then } dydz=rdrd\theta\\
&=\int_0^{2\pi}1d\theta\\
&=2\pi
\end{aligned} \tag{1.4}$$

故有，原式 $=1$

定理 1.2. Moment Generating Function

The m.g.f. of the distribution with p.d.f. given by Eq. (5.6.1) is

$$\psi(t)=e^{\mu t+\frac{1}{2}\sigma^2t^2} \tag{1.5}$$

for $-\infty<t<\infty$

♡

证明

$$\begin{aligned}
\psi(t)&=E(e^{tX})\\
&=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}e^{tx}dx\\
&=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{tx-\frac{(x-\mu)^2}{2\sigma^2}}dx
\end{aligned} \tag{1.6}$$

下面来分析 $tx-\frac{(x-\mu)^2}{2\sigma^2}$

$$
\begin{aligned}
tx - \frac{(x-\mu)^2}{2\sigma^2} &= -\frac{x^2-2\mu x+\mu^2-2t\sigma^2 x}{2\sigma^2} \\
&= -\frac{x^2-(2\mu+2t\sigma^2)x+\mu^2}{2\sigma^2} \\
&= -\frac{[x-(\mu+t\sigma^2)]^2+\mu^2-(\mu+t\sigma^2)^2}{2\sigma^2} \\
&= \mu t + \frac{1}{2}t^2\sigma^2 - \frac{[x-(\mu+t\sigma^2)]^2}{2\sigma^2}
\end{aligned}
\quad \text{合并成 } (x-\mu)^2 \text{ 的形式}, \ \text{简化} \tag{1.7}
$$

因此，原式 $\psi(t)$ 为

$$
\begin{aligned}
\psi(t) &= e^{\mu t+\frac{1}{2}t^2\sigma^2} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{[x-(\mu+t\sigma^2)]^2}{2\sigma^2}} \\
&= e^{\mu t+\frac{1}{2}t^2\sigma^2}
\end{aligned}
\tag{1.8}
$$

定理 1.3. Mean and Variance

The mean and variance of the distribution with p.d.f. given by Eq. (5.6.1) are and 2 , respectively. ♡

证明 $\psi(t)$ 的一阶导数和二阶导数为：

$$
\begin{aligned}
\psi^{'}(t) &= (\mu + t\sigma^2)e^{\mu t+\frac{1}{2}t^2\sigma^2} \\
\psi^{''}(t) &= ([\mu + t\sigma^2]^2 + \sigma^2)e^{\mu t+\frac{1}{2}t^2\sigma^2}
\end{aligned}
$$

在 $t = 0$ 处，

$$
\begin{aligned}
\psi^{'}(0) &= \mu \\
\psi^{''}(0) &= \mu^2 + \sigma^2
\end{aligned}
$$

因此，

$$
\begin{aligned}
E(X) &= \psi^{'}(0) = \mu \\
Var(X) &= \psi^{''}0 - [\psi^{'}(0)]^2 = \sigma^2
\end{aligned}
$$

定理 1.4. Linear Transformations

If X has the normal distribution with mean and variance 2 and if Y = aX + b, where a and b are given constants and a = 0, then Y has the normal distribution with mean a + b and variance a 2 2 . ♡

证明 已知 X 的 $m.g.f$ 为 $\psi(t) = e^{\mu t+\frac{1}{2}\sigma^2 t^2}$，令 ψ_Y 记作 Y 的 $m.g.f$，则有

$$
\begin{aligned}
\psi_Y(t) &= E(e^{t(aX+b)}) \\
&= e^{tb}E(e^{taX}) \\
&= e^{tb}\psi(at) \\
&= e^{tb}e^{a\mu t+\frac{1}{2}a^2\sigma^2 t^2} \\
&= e^{(a\mu+b)t+\frac{1}{2}a^2\sigma^2 t^2}
\end{aligned}
\tag{1.9}
$$

因此，均值为 $a\mu + b$，方差为 $a^2\sigma^2$

定义 1.2. Standard Normal Distribution

The normal distribution with mean 0 and variance 1 is called the **standard normal distribution**. The p.d.f. of the standard normal distribution is usually denoted by the symbol ϕ, and the c.d.f. is denoted by the symbol Φ. Thus,

$$
\phi(x) = f(x|0,1) = \frac{1}{\sqrt{2\pi e^{-\frac{x^2}{2}}}} \quad \text{for } -\infty < x < \infty \tag{1.10}
$$

and

$$\Phi(x) = \int_{-\infty}^{x} \phi(u)du \quad \text{for} -\infty < x < \infty \tag{1.11}$$

♣

定理 1.5. Consequences of Symmetry

For all x and all 0 < p < 1

$$\Phi(-x) = 1 - \Phi(x) \quad and \quad \Phi^{-1}(p) = -\Phi^{-1}(1-p) \tag{1.12}$$

♡

证明 由于 $\phi(x)$ 是关于 y 轴的偶函数。因此，对于所有的 $x(-\infty < x < \infty)$, $p(X \leqslant x) = p(X \geqslant x)$, 即 $\Phi(x) = 1 - \Phi(-x)$

第二个公式，$x = \Phi^{-1}(p)$，$-x = \Phi^{-1}(1-p)$

定理 1.6. Converting Normal Distributions to Standard

Let X have the normal distribution with mean and variance 2 . Let F be the c.d.f. of X. Then Z = (X)/ has the standard normal distribution, and, for all x and all 0 < p < 1,

$$F(x) = \Phi(\frac{x-\mu}{\sigma}) \tag{1.13}$$

$$F^{-1}(p) = \mu + \sigma\Phi^{-1}(p) \tag{1.14}$$

♡

证明 令 $Z = \frac{X-\mu}{\sigma}$,

$$F(x) = p(X \leqslant x) = p(\frac{X-\mu}{\sigma} \leqslant \frac{x-\mu}{\sigma}) = p(Z \leqslant \frac{x-\mu}{\sigma}) = \Phi(\frac{x-\mu}{\sigma}) \tag{1.15}$$

定理 1.7. Linear Combinations of Normally Distributed Variables

If the random variables $X_1, \ldots, X_k$ are independent and if X_i has the normal distribution with mean μ_i and variance $\sigma_i^2 (i = 1, \cdots, k)$, then the sum $X_1, \cdots, X_k$ has the normal distribution with mean $\mu_1, \cdots, \mu_k$ and variance $\sigma_1^2, \cdots, \sigma_k^2$.

♡

证明 已知，X_i 的 m.g.f 为 $\psi_i(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$，设 $X_1 + \cdots + X_k$ 的 m.g.f 为 $\psi(x)$。由于独立性，可得

$$\begin{aligned} \psi(t) &= \prod_{i=1}^{k} \psi_i(t) \\ &= e^{(\sum\limits_{i=1}^{k} \mu_i)t + \frac{1}{2}(\sum\limits_{i=1}^{k} \sigma_i^2)t^2} \end{aligned} \tag{1.16}$$

定义 1.3. Sample Mean

Let $X_1, \ldots, X_n$ be random variables. The average of these n random variables, $\frac{1}{n}\sum\limits_{i=1}^{n} X_i$, is called their sample mean and is commonly denoted $\bar{X}_n$.

♣

推论 1.1

Suppose that the random variables $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean μ and variance σ^2 , and let $\bar{X}_n$ denote their sample mean. Then $\bar{X}_n$ has the normal distribution with mean μ and variance $\frac{\sigma^2}{n}$.

♡

1.2 Gamma 分布

定义 1.4. The Gamma Function

For each positive number , let the value $\Gamma(\alpha)$ be dened by the following integral:

$$\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1}e^{-x}dx \tag{1.17}$$

The function Γ dened by Eq. (5.7.2) for $\alpha > 0$ is called the gamma function.

定理 1.8

if $\alpha > 1$, then

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) \tag{1.18}$$

证明 We shall apply the method of integration by parts to the integral in Eq. (5.7.2). If we let $u = x^{\alpha-1}$ and $dv = e^{-x}dx$, then $du = (\alpha-1)x^{\alpha-2}dx$ and $v = -e^{-x}$. Therefore,

$$\begin{aligned}\Gamma(\alpha) &= \int_0^\infty udv = [uv]_0^\infty - \int_0^\infty vdu \\ &= [-x^{\alpha-1}e^{-x}]_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx \\ &= 0 + (\alpha-1)\Gamma(\alpha-1)\end{aligned} \tag{1.19}$$

定理 1.9

For every positive integer n,

$$\Gamma(n) = (n-1)! \tag{1.20}$$

证明

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-1) \\ &= (n-1)(n-2)\Gamma(n-2) \\ &= (n-1)(n-2)\cdots(1)\Gamma(1) \\ &= (n-1)!\end{aligned} \tag{1.21}$$

定理 1.10

For each > 0 and each > 0,

$$\int_0^\infty x^{\alpha-1}e^{-\beta x}dx = \frac{\Gamma(\alpha)}{\beta^\alpha} \tag{1.22}$$

证明 令 $y = \beta x$，则有 $x = y/\beta$，以及 $dx = dy/\beta$。

$$\begin{aligned}\textstyle\int_0^\infty x^{\alpha-1}e^{-\beta x}dx &= \textstyle\int_0^\infty \frac{y^{\alpha-1}}{\beta^{\alpha-1}}e^{-y}\frac{1}{\beta}dy \\ &= \textstyle\frac{1}{\beta^\alpha}\int_0^\infty y^{\alpha-1}e^{-y}dy \\ &= \textstyle\frac{\Gamma(\alpha)}{\beta^\alpha}\end{aligned} \tag{1.23}$$

定义 1.5. Gamma Distributions

Let α and β be positive numbers. A random variable X has the gamma distribution with parameters α and β if X has a continuous distribution for which the $p.d.f.$ is

$$f(x|\alpha,\beta)=\begin{cases}\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & \text{for } x>0\\ 0 & \text{for } x\leqslant 0\end{cases} \tag{1.24}$$

♣

定理 1.11. Moments

Let X have the gamma distribution with parameters α and β. For $k=1,2,\ldots$

$$E(X^k)=\frac{\Gamma(\alpha+k)}{\beta^k\Gamma(\alpha)}=\frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k} \tag{1.25}$$

特别的，$E(X)=\frac{\alpha}{\beta}$，$Var(X)=\frac{\alpha}{\beta}$

♡

证明

$$\begin{aligned}E(X^k)&=\int_0^\infty x^k f(x|\alpha,\beta)dx=\int_0^\infty x^k\cdot\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\\&=\frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{\alpha+k-1}e^{-\beta x}dx\\&=\frac{\beta^\alpha}{\Gamma(\alpha)}\cdot\frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}}\\&=\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\beta^k}\end{aligned} \tag{1.26}$$

因此，$E(X)=\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\beta}=\frac{\alpha}{\beta}$，$Var(X)=E(X^2)-E^2(X)=\frac{(\alpha+1)\alpha}{\beta^2}-\frac{\alpha^2}{\beta^2}=\frac{\alpha}{\beta^2}$

定理 1.12. Moment Generating Function

Let X have the gamma distribution with parameters and . The m.g.f. of X is

$$\psi(t)=(\frac{\beta}{\beta-t})^\alpha \quad \text{for } t<\beta \tag{1.27}$$

♡

证明

$$\begin{aligned}\psi(t)=E(e^{tX})&=\int_0^\infty e^{tx}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx\\&=\frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{\alpha-1}e^{-(\beta-t)x}dx\\&=\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{(\Gamma(\alpha))}{(\beta-t)^\alpha}\\&=(\frac{\beta}{(\beta-t)})^\alpha\end{aligned} \tag{1.28}$$

定理 1.13

If the random variables X1, . . . , Xk are independent, and if Xi has the gamma distribution with parameters i and (i = 1, . . . , k), then the sum X1 + . . . + Xk has the gamma distribution with parameters 1 + . . . + k and .

♡

证明 If $\psi_i(t)$ denotes the m.g.f. of Xi , then it follows from Eq. (5.7.15) that for i = 1, . . . , k,

$$\psi(t)=\prod_{i=1}^k\psi_i(t)=(\frac{\beta}{\beta-t})^{\alpha_1+\cdots+\alpha_k} \quad \text{for } t<\beta \tag{1.29}$$

The m.g.f. can now be recognized as the m.g.f. of the gamma distribution with parameters 1 + . . . + k and . Hence, the sum X1 + . . . + Xk must have this gamma distribution.

1.3 指数分布

定义 1.6. Exponential Distributions

Let $\beta > 0$. A random variable X has the exponential distribution with parameter β if X has a continuous distribution with the $p.d.f.$

$$f(x|\beta) = \begin{cases} \beta e^{-\beta x} & \text{for } x > 0 \\ 0 & \text{for } x \leqslant 0 \end{cases} \tag{1.30}$$

♣

定理 1.14

The exponential distribution with parameter β is the same as the gamma distribution with parameters $\alpha = 1$ and β. If X has the exponential distribution with parameter β, then $E(X) = \frac{1}{\beta}$ and $Var(X) = \frac{1}{\beta^2}$ and the m.g.f. of X is

$$\psi(t) = \frac{\beta}{\beta - t} \quad \text{for } t < \beta \tag{1.31}$$

♡

证明 根据 Gamma 分布，指数分布是 Gamma 的一个特例，$Gamma(\alpha = 1, \beta)$，因此期望 $E(X) = \frac{\alpha}{\beta} = \frac{1}{\beta}$，$Var(X) = \frac{\alpha}{\beta^2} = \frac{1}{\beta^2}$，$\psi(t) = (\frac{\beta}{\beta - t})^\alpha = \frac{\beta}{\beta - t}$

定理 1.15. Memoryless Property of Exponential Distributions

Let X have the exponential distribution with parameter β, and let $t > 0$. Then for every number $h > 0$,

$$p(X \geqslant t + h|X \geqslant t) = p(X \geqslant h) \tag{1.32}$$

♡

证明 for each $t > 0$,

$$p(X \geqslant t) = \int_t^\infty \beta e^{-\beta x} dx = e^{-\beta t} \tag{1.33}$$

因此，对于所有的 $t > 0$ 以及 $h > 0$，

$$\begin{aligned} p(X \geqslant t + h|x \geqslant t) &= \frac{p(x \geqslant t + h)}{p(X \geqslant t)} \\ &= \frac{e^{-\beta(t+h)}}{e^{-\beta t}} \\ &= e^{-\beta h} \\ &= p(X \geqslant h) \end{aligned} \tag{1.34}$$

定理 1.16

Suppose that the variables $X_1, \ldots, X_n$ form a random sample from the exponential distribution with parameter β. Then the distribution of $Y_1 = \min\{X_1, \ldots, X_n\}$ will be the exponential distribution with parameter $n\beta$.

♡

证明 for every number $t > 0$,

$$\begin{aligned} p(Y_1 > t) &= p(X_1 > t, \dots, X_n > t) \\ &= p(X_1 > t) \cdots p(X_n > t) \\ &= e^{-\beta t} \cdots e^{-\beta t} \\ &= e^{-n\beta t} \end{aligned} \tag{1.35}$$

1.4 Beta 分布

定义 1.7. The Beta Function

for each positive α and β, define

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \tag{1.36}$$

the function B is called the *beta function*.

定理 1.17

for all $\alpha, \beta > 0$,

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{1.37}$$

定义 1.8. Beta Distributions

Let $\alpha, \beta > 0$ and let X be a random variable with p.d.f

$$f(x|\alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1.38}$$

定理 1.18. Moments

Suppose that X has the beta distribution with parameters α and β. Then for each positive integer k,

$$E(X^k) = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+k-1)} \tag{1.39}$$

特别地，

$$E(X) = \frac{\alpha}{\alpha+\beta} \tag{1.40}$$

$$Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{1.41}$$

证明 for $k = 1, 2, \dots$

$$\begin{aligned} E(X^k) &= \int_0^1 x^k f(x|\alpha, \beta)dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha+k-1}(1-x)^{\beta-1}dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+k+\beta)} \end{aligned} \tag{1.42}$$

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

第 2 章 large random samples

2.1 大数定律

定理 2.1. Markov Inequality

Suppose that X is a random variable such that $p(X \geqslant 0) = 1$. Then for every real number $t > 0$,

$$p(X \geqslant t) \leqslant \frac{E(X)}{t} \tag{2.1}$$

证明 离散情况下

$$E(X) = \sum_{x} xf(x) = \sum_{x<t} xf(x) + \sum_{x \geqslant t} xf(x) \tag{2.2}$$

由于 $X \geqslant 0$，所有项都大于 0。因此

$$E(X) \geqslant \sum_{x \geqslant t} xf(x) \geqslant \sum_{x \geqslant t} tf(x) = t \cdot p(X \geqslant t) \tag{2.3}$$

连续情况下

$$\begin{aligned} E(X) &= \int_0^{\infty} xf(x)dx \\ &= \int_0^{t} xf(x)dx + \int_t^{\infty} xf(x)dx \\ &\geqslant \int_t^{\infty} xf(x)dx \\ &\geqslant \int_t^{\infty} tf(x)dx \\ &= t \cdot p(X \geqslant t) \end{aligned} \tag{2.4}$$

定理 2.2. Chebyshev Inequality

Let X be a random variable for which $Var(X)$ exists. Then for every number $t > 0$,

$$p(|X - E(X)| \geqslant t) \leqslant \frac{Var(X)}{t^2} \tag{2.5}$$

证明 令 $Y = [X - E(X)^2]$，则 $E(Y) = Var(X)$，

$$p(|X - E(X)| \geqslant t) = p(Y \geqslant t^2) \leqslant \frac{E(Y)}{t^2} = \frac{Var(X)}{t^2} \tag{2.6}$$

定理 2.3. Mean and Variance of the Sample Mean

Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ and variance σ^2 . Let $\bar{X}_n$ be the sample mean. Then $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \frac{\sigma^2}{n}$.

定义 2.1. Convergence in Probability

A sequence $Z_1, Z_2, \ldots$ of random variables converges to b in probability if for every number $\varepsilon > 0$,

$$\lim_{n \to \infty} p(|Z_n - b| < \varepsilon) = 1 \tag{2.7}$$

This property is denoted by

$$Z_n \xrightarrow{p} b \tag{2.8}$$

and is sometimes stated simply as Z_n converges to b in probability.

定理 2.4. Law of Large Numbers

Suppose that $X_1, ..., X_n$ form a random sample from a distribution for which the mean is μ and for which the variance is nite. Let $\bar{X}_n$ denote the sample mean. Then

$$\bar{X}_n \xrightarrow{p} \mu \tag{2.9}$$

证明 Let the variance of each X_i be σ^2. It then follows from the Chebyshev inequality that for every number $\varepsilon > 0$,

$$p(|\bar{X}_n - \mu| < \varepsilon) \geqslant 1 - \frac{\sigma^2}{n\varepsilon^2} \tag{2.10}$$

因此，

$$\lim_{n\to\infty} p(|\bar{X}_n - \mu| < \varepsilon) = 1 \tag{2.11}$$

which means that $\bar{X}_n \xrightarrow{p} \mu$.

定理 2.5. Continuous Functions of Random Variables

If $Z_n \xrightarrow{p} b$, and if $g(z)$ is a function that is continuous at $z = b$, then$g(Z_n) \xrightarrow{p} g(b)$.