目录

序

第 1 章　特殊分布

1.1　泊松分布

the family of Poisson distributions is used to model the number of such arrivals that occur in a **fixed time period**.

定义 1.1. Poisson Distribution 泊松分布

oisson Distribution. Let > 0. A random variable X has the Poisson distribution with mean λ if the p.f. of X is as follows:

$$f(x|\lambda) = \begin{cases} \frac{\lambda^x}{x!}e^{-\lambda} & \text{for x in 0, 1, ...} \\ 0 & \text{otherwise} \end{cases} \tag{1.1}$$

♣

定理 1.1. Poisson Mean

The mean of the distribution with p.f. equal to λ.

♡

证明

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x f(x|\lambda) \\ &= \sum_{x=1}^{\infty} x f(x|\lambda) \\ &= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \qquad \textit{let } y = x-1 \\ &= \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} \\ &= \lambda \end{aligned} \tag{1.2}$$

定理 1.2. Poisson Variance

The variance of the Poisson distribution with mean λ is also λ.

♡

证明

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1) f(x|\lambda) \\ &= \sum_{x=2}^{\infty} x(x-1) f(x|\lambda) \\ &= \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} \qquad \textit{let } y = x-2 \\ &= \lambda^2 \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} \\ &= \lambda^2 \end{aligned} \tag{1.3}$$

$$E(X^2) - E(X) = \lambda^2$$

因此，

$$
\begin{aligned}
Var(X) &= E(X^2) - E^2(X) \\
&= \lambda^2 + E(x) - E^2(X) \\
&= \lambda
\end{aligned} \tag{1.4}
$$

定理 1.3. Poisson Moment Generating Function

The m.g.f. of the Poisson distribution with mean λ is

$$\psi(t) = e^{\lambda(e^t-1)} \tag{1.5}$$

♡

证明 对于所有的 $t(-\infty < t < \infty)$，

$$
\begin{aligned}
\psi(t) &= E(e^{tX}) \\
&= \sum_{x=0}^{\infty} \frac{e^{tx}e^{-\lambda}\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{((\lambda e^t))^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t} \\
&= e^{\lambda(e^t-1)}
\end{aligned} \tag{1.6}
$$

定理 1.4

If the random variables $X_1, \cdots, X_k$ are independent and if X_i has the Poisson distribution with mean $\lambda_i(i = 1, ..., k)$, then the sum $X_1 + \cdots + X_k$ has the Poisson distribution with mean $\lambda_1 + \cdots + \lambda_k$.

♡

证明 let $\psi_i(t)$ 记为 X_i 的概率密度函数，$i = 1, \cdots, k$，令 $\psi(t)$ 为 $X_1 + \cdots + X_k$ 的概率密度函数，因为 $X_1, \cdots, X_k$ 是独立的，因此

$$\psi(t) = \prod_{i=1}^{k} \psi(t) = \prod_{i=1}^{k} e^{\lambda_i(e^t-1)} = e^{(\lambda_1+\cdots+\lambda_k)(e^t-1)} \tag{1.7}$$

定理 1.5. Closeness of Binomial and Poisson Distributions

For each integer n and each $0 < p < 1$, let $f(x|n, p)$ denote the p.f. of the binomial distribution with parameters n and p. Let $f(x|\lambda)$ denote the p.f. of the Poisson distribution with mean λ. Let $\{p_n\}_{n=1}^{\infty}$ be a sequence of numbers between 0 and 1 such that $\lim_{n\to\infty} np_n = \lambda$. Then

$$\lim_{n\to\infty} f(x|n, p_n) = f(x|\lambda)$$

for all $x = 0, 1, \ldots$

♡

证明

$$
\begin{aligned}
f(x|n, p_n) &= \tfrac{n(n-1)\cdots(n-x+1)}{x} p_n^x (1-p_n)^{n-x} \\
&= \tfrac{\lambda_n^x}{x!} \cdot \tfrac{n}{n} \cdot \tfrac{n-1}{n} \cdots \tfrac{n-x+1}{n} (1 - \tfrac{\lambda_n}{n})^n (1 - \tfrac{\lambda_n}{n})^{-x}
\end{aligned}
\quad
\begin{array}{l}
\textit{let } \lambda_n = nP_n, \\
\textit{so that } \lim\limits_{n\to\infty} \lambda_n = \lambda
\end{array} \tag{1.8}
$$

对于所有的 $x \geq 0$，

$$\lim_{n\to\infty} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-x+1}{n} \cdot (1 - \frac{\lambda_n}{n})^{-x} = 1 \tag{1.9}$$

定义 1.2. Poisson Process 泊松过程

Poisson process with rate λ per unit time is a process that satises the following two properties:

1. The number of arrivals in every fixed interval of time of length t has the Poisson distribution with mean λt.
2. The numbers of arrivals in every collection of disjoint time intervals are independent.

1.2 正态分布

定义 1.3. 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 $p.d.f.$

$$f(x|\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1.10}$$

for $-\infty<x<\infty$

定理 1.6

正态分布的概率密度函数积分为 1。

$$\int_{-\infty}^{\infty} f(x|\mu,\sigma^2)dx=1 \tag{1.11}$$

证明 let $y=\frac{x-\mu}{\sigma}$，then

$$\int_{-\infty}^{\infty} f(x|\mu,\sigma^2)dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy \tag{1.12}$$

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

$$\begin{aligned}
I^2 &= \int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{y^2+z^2}{2}}dydz \qquad \textit{let } y=r\cos\theta,\ z=r\sin\theta \\
&= \int_{0}^{2\pi}\int_{0}^{\infty}e^{-\frac{r^2}{2}}rdrd\theta \qquad \textit{then } dydz=rdrd\theta \\
&= \int_{0}^{2\pi}1d\theta \\
&= 2\pi
\end{aligned} \tag{1.13}$$

故有，原式 $=1$

定理 1.7. Moment Generating Function

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

$$\psi(t)=e^{\mu t+\frac{1}{2}\sigma^2t^2} \tag{1.14}$$

for $-\infty<t<\infty$

证明

$$\begin{aligned}
\psi(t) &= E(e^{tX}) \\
&= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}e^{tx}dx \\
&= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{tx-\frac{(x-\mu)^2}{2\sigma^2}}dx
\end{aligned} \tag{1.15}$$

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

$$
\begin{aligned}
tx - \frac{(x-\mu)^2}{2\sigma^2} &= -\frac{x^2-2\mu x+\mu^2-2t\sigma^2 x}{2\sigma^2} \\
&= -\frac{x^2-(2\mu+2t\sigma^2)x+\mu^2}{2\sigma^2} \\
&= -\frac{[x-(\mu+t\sigma^2)]^2+\mu^2-(\mu+t\sigma^2)^2}{2\sigma^2} && \text{合并成 } (x-\mu)^2 \text{ 的形式} \\
&= \mu t + \frac{1}{2}t^2\sigma^2 - \frac{[x-(\mu+t\sigma^2)]^2}{2\sigma^2} && \text{简化}
\end{aligned} \tag{1.16}
$$

因此，原式 $\psi(t)$ 为

$$
\begin{aligned}
\psi(t) &= e^{\mu t+\frac{1}{2}t^2\sigma^2} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{[x-(\mu+t\sigma^2)]^2}{2\sigma^2}} \\
&= e^{\mu t+\frac{1}{2}t^2\sigma^2}
\end{aligned} \tag{1.17}
$$

定理 1.8. 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 1.9. 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

$$
\begin{aligned}
\psi_Y(t) &= E(e^{t(aX+b)}) \\
&= e^{tb}E(e^{taX}) \\
&= e^{tb}\psi(at) \\
&= e^{tb}e^{a\mu t+\frac{1}{2}a^2\sigma^2 t^2} \\
&= e^{(a\mu+b)t+\frac{1}{2}a^2\sigma^2 t^2}
\end{aligned} \tag{1.18}
$$

因此，均值为 $a\mu + b$，方差为 $a^2\sigma^2$

定义 1.4. 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,

$$
\phi(x) = f(x|0,1) = \frac{1}{\sqrt{2\pi}e^{-\frac{x^2}{2}}} \quad \text{for } -\infty < x < \infty \tag{1.19}
$$

and

$$\Phi(x)=\int_{-\infty}^{x}\phi(u)du \quad \text{for} -\infty<x<\infty \tag{1.20}$$

定理 1.10. Consequences of Symmetry

For all x and all 0 < p < 1

$$\Phi(-x)=1-\Phi(x) \quad and \quad \Phi^{-1}(p)=-\Phi^{-1}(1-p) \tag{1.21}$$

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 1.11. 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,

$$F(x)=\Phi(\frac{x-\mu}{\sigma}) \tag{1.22}$$

$$F^{-1}(p)=\mu+\sigma\Phi^{-1}(p) \tag{1.23}$$

证明 令 $Z=\frac{X-\mu}{\sigma}$，

$$F(x)=p(X\leqslant x)=p(\frac{X-\mu}{\sigma}\leqslant\frac{x-\mu}{\sigma})=p(Z\leqslant\frac{x-\mu}{\sigma})=\Phi(\frac{x-\mu}{\sigma}) \tag{1.24}$$

定理 1.12. 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 $\psi_i(t)=e^{\mu t+\frac{1}{2}\sigma^2t^2}$，设 $X_1+\cdots+X_k$ 的 m.g.f 为 $\psi(x)$。由于独立性，可得

$$\begin{aligned}\psi(t)&=\prod_{i=1}^{k}\psi_i(t)\\&=e^{(\sum\limits_{i=1}^{k}\mu_i)t+\frac{1}{2}(\sum\limits_{i=1}^{k}\sigma_i^2)t^2}\end{aligned} \tag{1.25}$$

定义 1.5. 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 $\frac{\sigma^2}{n}$.

1.3 Gamma 分布

定义 1.6. The Gamma Function

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

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx \tag{1.26}$$

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

定理 1.13

if $\alpha > 1$, then

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) \tag{1.27}$$

♡

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,

$$\begin{aligned}\Gamma(\alpha) &= \int_0^\infty udv = [uv]_0^\infty - \int_0^\infty vdu \\ &= [-x^{\alpha-1}e^{-x}]_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx \\ &= 0 + (\alpha-1)\Gamma(\alpha-1)\end{aligned} \tag{1.28}$$

定理 1.14

For every positive integer n,

$$\Gamma(n) = (n-1)! \tag{1.29}$$

证明

$$\begin{aligned}\Gamma(n) &= (n-1)\Gamma(n-1) \\ &= (n-1)(n-2)\Gamma(n-2) \\ &= (n-1)(n-2)\cdots(1)\Gamma(1) \\ &= (n-1)!\end{aligned} \tag{1.30}$$

♡

定理 1.15

For each > 0 and each > 0,

$$\int_0^\infty x^{\alpha-1}e^{-\beta x}dx = \frac{\Gamma(\alpha)}{\beta^\alpha} \tag{1.31}$$

♡

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

$$\begin{aligned}\int_0^\infty x^{\alpha-1}e^{-\beta x}dx &= \int_0^\infty \frac{y^{\alpha-1}}{\beta^{\alpha-1}}e^{-y}\frac{1}{\beta}dy \\ &= \frac{1}{\beta^\alpha}\int_0^\infty y^{\alpha-1}e^{-y}dy \\ &= \frac{\Gamma(\alpha)}{\beta^\alpha}\end{aligned} \tag{1.32}$$

定义 1.7. 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

$$f(x|\alpha,\beta)=\begin{cases}\dfrac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & \text{for } x>0\\ 0 & \text{for } x\leqslant 0\end{cases} \tag{1.33}$$

♣

定理 1.16. Moments

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

$$E(X^k)=\frac{\Gamma(\alpha+k)}{\beta^k\Gamma(\alpha)}=\frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k} \tag{1.34}$$

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

♡

证明

$$\begin{aligned}E(X^k)&=\int_0^{\infty}x^kf(x|\alpha,\beta)dx=\int_0^{\infty}x^k\cdot\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\\&=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty}x^{\alpha+k-1}e^{-\beta x}dx\\&=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\cdot\frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}}\\&=\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\beta^k}\end{aligned} \tag{1.35}$$

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 1.17. Moment Generating Function

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

$$\psi(t)=(\frac{\beta}{\beta-t})^{\alpha}\quad\text{for } t<\beta \tag{1.36}$$

♡

证明

$$\begin{aligned}\psi(t)=E(e^{tX})&=\int_0^{\infty}e^{tx}\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx\\&=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty}x^{\alpha-1}e^{-(\beta-t)x}dx\\&=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\frac{(\Gamma(\alpha))}{(\beta-t)^{\alpha}}\\&=(\frac{\beta}{(\beta-t)})^{\alpha}\end{aligned} \tag{1.37}$$

定理 1.18

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,

$$\psi(t)=\prod_{i=1}^{k}\psi_i(t)=(\frac{\beta}{\beta-t})^{\alpha_1+\cdots+\alpha_k}\quad\text{for } t<\beta \tag{1.38}$$

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.

1.4 指数分布

定义 1.8. Exponential Distributions

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

$$f(x|\beta) = \begin{cases} \beta e^{-\beta x} & \text{for } x > 0 \\ 0 & \text{for } x \leqslant 0 \end{cases} \tag{1.39}$$

♣

定理 1.19

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

$$\psi(t) = \frac{\beta}{\beta - t} \quad \text{for } t < \beta \tag{1.40}$$

♡

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 1.20. Memoryless Property of Exponential Distributions

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

$$p(X \geqslant t + h|X \geqslant t) = p(X \geqslant h) \tag{1.41}$$

♡

证明 for each $t > 0$,

$$p(X \geqslant t) = \int_t^\infty \beta e^{-\beta x} dx = e^{-\beta t} \tag{1.42}$$

因此，对于所有的 $t > 0$ 以及 $h > 0$，

$$\begin{aligned} p(X \geqslant t + h|x \geqslant t) &= \frac{p(x \geqslant t + h)}{p(X \geqslant t)} \\ &= \frac{e^{-\beta(t+h)}}{e^{-\beta t}} \\ &= e^{-\beta h} \\ &= p(X \geqslant h) \end{aligned} \tag{1.43}$$

定理 1.21

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 $t > 0$,

$$\begin{aligned} p(Y_1 > t) &= p(X_1 > t, \ldots, X_n > t) \\ &= p(X_1 > t) \cdots p(X_n > t) \\ &= e^{-\beta t} \cdots e^{-\beta t} \\ &= e^{-n\beta t} \end{aligned} \tag{1.44}$$

1.5 Beta 分布

定义 1.9. The Beta Function

for each positive α and β, define

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \tag{1.45}$$

the function B is called the *beta function*.

♣

定理 1.22

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

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{1.46}$$

♡

定义 1.10. Beta Distributions

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

$$f(x|\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1.47}$$

♣

定理 1.23. Moments

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

$$E(X^k) = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+k-1)} \tag{1.48}$$

特别地，

$$E(X) = \frac{\alpha}{\alpha+\beta} \tag{1.49}$$

$$Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{1.50}$$

♡

证明 for $k = 1, 2, \ldots$

$$\begin{aligned} E(X^k) &= \int_0^1 x^k f(x|\alpha,\beta)dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha+k-1}(1-x)^{\beta-1}dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+k+\beta)} \end{aligned} \tag{1.51}$$

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 大样本抽样

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 $X_1, \ldots, X_n$ 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$g(Z_n) \xrightarrow{\ p\ } g(b)$. ♡

2.2 中心极限定理

中心极限定理部分。

第 3 章 估计

3.1 statistical inference

统计推断

第 4 章 抽样分布

4.1 χ^2 分布

定义 4.1. χ^2 分布

For each positive number m, the gamma distribution with parameters $\alpha = \frac{m}{2}$, and $\beta = \frac{1}{2}$ is called the χ^2 distribution with m degrees of freedom.

If a random variable X has the χ^2 distribution with m degrees of freedom, the p.d.f. of X for $x > 0$ is

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} = \frac{1}{2^{m/2}\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} \tag{4.1}$$

Also, $f(x) = 0$ for $x \leqslant 0$.

♣

定理 4.1. Mean and Variance

f a random variable X has the χ^2 distribution with m degrees of freedom, then E(X) = m and Var(X) = 2m.

♡

m.g.f. of X is

$$\psi(t) = (\frac{1}{1-2t})^{m/2} \quad \text{for } t < \frac{1}{2} \tag{4.2}$$

定理 4.2

If the random variables $X_1, \ldots, X_k$ are independent and if X_i has the χ^2 distribution with m_i degrees of freedom $(i = 1, \ldots, k)$, then the sum $X_1 + \cdots + X_k$ has the χ^2 distribution with $m_1, \ldots, m_k$ degrees of freedom.

♡

证明 证明可用 Γ 分布

定理 4.3

Let X have the standard normal distribution. Then the random variable $Y = X^2$ has the χ^2 distribution with one degree of freedom.

♡

证明 需要证明 $Y \sim \chi(1)$，或者说，服从 $Y \sim Gamma(\frac{1}{2}, \frac{1}{2})$

Let $f(y)$ and $F(y)$ denote, respectively, the p.d.f. and the c.d.f. of Y . Also, since X has the standard normal distribution, we shall let $\psi(x)$ and $\Psi(x)$ denote the p.d.f. and the c.d.f. of X. Then for y > 0,

$$\begin{aligned} p(Y \leqslant y) &= p(X^2 \leqslant y) \\ &= p(-y^{1/2} \leqslant X \leqslant y^{1/2}) \\ &= \Phi(y^{1/2}) - \Phi(-y^{1/2}) \end{aligned} \tag{4.3}$$

由于 $f(y) = F^{'}(y)$ 以及 $\phi(x) = \Phi^{'}(x)$，因此

$$f(y) = \Phi^{'}(y^{1/2})(\frac{1}{2}y^{-1/2}) - \Phi^{'}(-y^{1/2})(-\frac{1}{2}y^{-1/2}) \tag{4.4}$$

同时，因为 $\phi(y^{1/2}) = \phi(-y^{1/2}) = \frac{1}{\sqrt{2\pi}} e^{-y/2}$，$\Gamma(\frac{1}{2}) = \pi^{1/2}$ 此时，

$$\begin{aligned} f(y) &= \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} \\ &\sim Gamma(\frac{1}{2}, \frac{1}{2}) \\ &\sim \chi^2(1) \end{aligned} \tag{4.5}$$

推论 4.1. I

the random variables $X_1, \ldots, X_m$ are i.i.d. with the standard normal distribution, then the sum of squares $X_1^2 + \cdots + X_m^2$ 服从 χ^2 分布，自由度为 m。

If the random variables X1, . . . , Xm are i.i.d. with the standard normal distribution,2 has the 2 distribution with m degrees of then the sum of squares X12 + . . . + Xm freedom.

4.2 t Distributions

定义 4.2. t Distributions

Consider two independent random variables Y and Z, such that Y has the 2 distribution with m degrees of freedom and Z has the standard normal distribution. Suppose that a random variable X is dened by the equation

$$X = \frac{Z}{(\frac{Y}{m})^{1/2}} \tag{4.6}$$

Then the distribution of X is called the t distribution with m degrees of freedom.

定理 4.4. Probability Density Function

t 分布在 $m\frac{a}{b}$ 自由度下的概率密度函数为：

$$\frac{\Gamma(\frac{m+1}{2})}{(m\pi)^{1/2}\Gamma(\frac{m}{2})} \left(1 + \frac{x^2}{m}\right)^{-(m+1)/2} \quad \text{for } -\infty < x < \infty \tag{4.7}$$

定理 4.5

Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean μ and variance σ^2 . Let $\bar{X}_n$ denote the sample mean, and dene

$$\sigma' = \left[\frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}{n-1} \right]^{1/2} \tag{4.8}$$

则 $\frac{n^{1/2}(\bar{X}_n - \mu)}{\sigma'} \sim t(n-1)$

证明 定义 $S_n^2 = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$，定义 $Z = \frac{\bar{X}_n - \mu}{\sigma/n}$ 以及 $Y = \frac{S_n^2}{\sigma^2}$

由于 Y 与 Z 独立，且 $Y \sim \chi^2(n-1)$，$Z \sim N(0,1)$，定义 U 为

$$U = \frac{Z}{\left(\frac{Y}{n-1}\right)^{1/2}} \\ = \tag{4.9}$$

4.3 置信区间

定义 4.3. Condence Interval

Let $X = (X_1, \ldots, X_n)$ be a random sample from a distribution that depends on a parameter (or parameter vector) . Let $g()$ be a real-valued function of . Let $A \leqslant B$ be two statistics that have the property that for all values of ,

$$p(A < g(\theta) < B) \geqslant \gamma \tag{4.10}$$

Then the random interval (A, B) is called a coefficient γ condence interval for $g()$ or a 100γ percent condence interval for $g()$. If the inequality "$\geqslant \gamma$" in Eq. (8.5.4) is an equality for all θ , the condence interval is called exact. After the values of the random variables $X_1, \ldots, X_n$ in the random sample have been observed, the values of $A = a$ and $B = b$ are computed, and the interval (a, b) is called the observed value of the condence interval. ♣

定义 4.4. One-Sided Condence Intervals/Limits

Let $X = (X1, ..., Xn)$ be a random sample from a distribution that depends on a parameter (or parameter vector) θ. Let $g(\theta)$ be a real-valued function of θ . Let A be a statistic that has the property that for all values of θ,

$$p(A < g(\theta)) \geqslant \gamma \tag{4.11}$$

. Then the random interval (A, ∞) is called a one-sided coeffcient γ condence interval for $g(\theta)$ or a one-sided 100γ percent condence interval for $g(\theta)$. Also, A is called a coefcient γ lower condence limit for $g(\theta)$ or a 100γ percent lower condence limit for $g(\theta)$. Similarly, if B is a statistic such that

$$p(g(\theta) < B) \geqslant \theta \tag{4.12}$$

then (∞, B) is a one-sided coefcient γ condence interval for $g(\theta)$ or a one-sided 100γ percent condence interval for $g(\theta)$ and B is a coefcient g upper condence limit Condence Intervals for $g(\theta)$ or a 100γ percent upper condence limit for $g(\theta)$. If the inequality "$\geqslant \gamma$ " in either Eq. (8.5.5) or Eq. (8.5.6) is equality for all θ, the corresponding condence interval and condence limit are called exact. ♣

定理 4.6. One-Sided Condence Intervals for the Mean of a Normal Distribution

Let $X_1, \ldots, X_n$ be a random sample from the normal distribution with mean μ and variance. For each $0 < \gamma < 1$, the following statistics are, respectively, exact lower and upper coefcient γ condence limits for μ:

$$
\begin{aligned}
A &= \bar{X}_n - T_{n-1}^{-1}(\gamma)\frac{\sigma^{'}}{n^{1/2}} \\
B &= \bar{X}_n + T_{n-1}^{-1}(\gamma)\frac{\sigma^{'}}{n^{1/2}}
\end{aligned}
\tag{4.13}
$$

定义 4.5. p-value

In general, the p-value is the smallest level α_0 such that we would reject the null-hypothesis at level α_0 with the observed data.

An experimenter who rejects a null hypothesis if and only if the p-value is at most α_0 is using a test with level of signicance α_0. Similarly, an experimenter who wants a level α_0 test will reject the null hypothesis if and only if the p-value is at most α_0. For this reason, the p-value is sometimes called *the observed level of signicance*.

第 5 章 假设检验

5.1 Problems of Testing Hypotheses 假设检验问题

定义 5.1. Null and Alternative Hypotheses/Reject.

The hypothesis H_0 is called the null hypothesis and the hypothesis H_1 is called the alternative hypothesis. When performing a test, if we decide that θ lies in Ω_1, we are said to reject H_0. If we decide that θ lies in Ω_0, we are said not to reject Ω_0. ♣

定义 5.2. power function

Let δ be a test procedure. The function $\pi(\theta|\delta)$ is called the power function of the test δ. If S_1 denotes the critical region of δ, then the power function $\pi(\theta|\delta)$ is determined by the relation

$$\pi(\theta|\delta) = p(X \in S_1|\theta) \quad \text{for } \theta \in \Omega \tag{5.1}$$ ♣

定义 5.3. Type I/II Error

An erroneous decision to reject a true null hypothesis is a *type I error*, or an error of the rst kind. An erroneous decision not to reject a false null hypothesis is called a *type II error*, or an error of the second kind.

Null Hypothesis	True	False
reject	**I** 类错误	
not reject		**II** 类错误

♣

两类错误无法同时避免，我们尽可能的避免第一类错误（这种错误更严重）。即，使拒绝 H_0 出错的概率最低。或者，在一些情况下，有要证明的理论时，把要证明的结果放在备择假设 (H_1) 中。

定义 5.4. Level/Size

A test that satises (9.1.6) is called a level α_0 test, and we say that the test has level of signicance α_0. In addition, the size $\alpha(\delta)$ of a test δ is dened as follows:

$$\alpha(\delta) = \sup_{\theta \in \Omega_0} \pi(\theta|\delta) \tag{5.2}$$ ♣

推论 5.1

A test δ is a level α_0 test if and only if its size is at most α_0 (i.e., $\alpha(\delta) \leqslant \alpha_0$). If the null hypothesis is simple, that is, $H_0 : \theta = \theta_0$, then the size of δ will be $\alpha(\delta) = \pi(\theta_0|\delta)$. ♡

定理 5.1. Level and Unbiasedness of t Tests

令 $X = (X_1, \ldots, X_n)$ 来自于均值为 μ，方差为 σ^2 的正态分布。令 U 为统计量

$$U = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{s} \tag{5.3}$$

其中，$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$，令 c 为 $t(n-1)$ 分布的 $1-\alpha_0$ 分位数，令 δ 为检验过

程：如果 $U \geqslant c$ 则拒绝 H_0。则 power function $\pi(\mu,\sigma^2|\delta)$ 有以下性质：

1. $\pi(\mu,\sigma^2|\delta) = \alpha_0$ when $\mu = \mu_0$,
2. $\pi(\mu,\sigma^2|\delta) < \alpha_0$ when $\mu < \mu_0$,
3. $\pi(\mu,\sigma^2|\delta) > \alpha_0$ when $\mu > \mu_0$,
4. $\pi(\mu,\sigma^2|\delta) \to 0$ when $\mu \to -\infty$,
5. $\pi(\mu,\sigma^2|\delta) \to 1$ when $\mu \to \infty$,

另外，检验过程 δ 的 size 为 α_0 并且无偏。

证明 如果 $\mu = \mu_0$，则 U 服从 t 分布，自由度为 $n-1$，因此，

$$\pi(\mu_0,\sigma^2|\delta) = p(U \geqslant c|\mu_0,\sigma^2) = \alpha_0. \tag{5.4}$$

以上，证明了1.，对于2.和3.，定义：

$$U^* = \frac{n^{1/2}(\bar{X}_n - \mu)}{s} \quad \text{and} \quad W = \frac{n^{1/2}(\mu_0 - \mu)}{s} \tag{5.5}$$

构建 $U = U^* - W$，首先，假设 $\mu < \mu_0$，所以 $W > 0$，以下有：

$$\begin{aligned}\pi(\mu,\sigma^2|\delta) &= p(U \geqslant c|\mu,\sigma^2)\\ &= p(U^* - W \geqslant c|\mu,\sigma^2)\\ &= p(U^* \geqslant c + W|\mu,\sigma^2) && \text{由于 } W > 0\\ &< p(U^* \geqslant c|\mu,\sigma^2) && U^* \sim t(n-1)\\ &= \alpha_0\end{aligned} \tag{5.6}$$

同理，当 $\mu > \mu_0$ 时，即 $W < 0$，所以可推导出 $\pi(\mu,\sigma^2|\delta) > \alpha_0$when$\mu > \mu_0$

推论 5.2. t Tests for Hpotheses

$X = (X_1,\ldots,X_n)$ 来自于均值为 μ，方差为 σ^2 的正态分布。令 U 为统计量

$$U = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{s} \tag{5.7}$$

其中，$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$，令 c 为 $t(n-1)$ 分布的 $1-\alpha_0$ 分位数，令 δ 为检验过程：如果 $U \leqslant c$ 则拒绝 H_0。则 power function $\pi(\mu,\sigma^2|\delta)$ 有以下性质：

1. $\pi(\mu,\sigma^2|\delta) = \alpha_0$ when $\mu = \mu_0$,
2. $\pi(\mu,\sigma^2|\delta) > \alpha_0$ when $\mu < \mu_0$,
3. $\pi(\mu,\sigma^2|\delta) < \alpha_0$ when $\mu > \mu_0$,
4. $\pi(\mu,\sigma^2|\delta) \to 1$ when $\mu \to -\infty$,
5. $\pi(\mu,\sigma^2|\delta) \to 0$ when $\mu \to \infty$,

另外，检验过程 δ 的 size 为 α_0 并且无偏。

定理 5.2. p-value for t test

假设我们在检验一个假设，假设为 $\mu \geqslant \mu_0$ 或 $\mu \leqslant \mu_0$，令 $U = n^{1/2} \cdot \frac{\bar{X}_n - \mu_0}{s}$，令 u 为 U 的观测值。令 $T_{n-1}(\cdot)$ 为 $n-1$ 个自由的 t 分布的 $c.d.f.$，假设（$H_0 : \mu \leqslant \mu_0$）的 $p-value$ 为 $1 - T_{n-1}(u)$，同时假设（$H_0 : \mu \geqslant \mu_0$）的 $p-value$ 为 $T_{n-1}(u)$。

证明 令 $T_{n-1}^{-1}(\cdot)$ 表示 $n-1$ 个自由度的 t 分布的分位数。这是 T_{n-1} 是严格的增函数。拒绝原假设 $H_0 : \mu \leqslant \mu_0$ 当且仅当 $u \geqslant T_{n-1}^{-1}(1-\alpha_0)$，等价于 $T_{n-1}(u) \geqslant 1-\alpha_0$，等价于 $\alpha_0 \geqslant 1 - T_{n-1}(u)$