抽样分布

正态总体下样本均值和样本方差的分布

独立的正态随机变量的线性组合仍然服从正态分布:

Thr

定理 1. 设随机变量 $X_1, \cdots, X_n$ 相互独立, 且 $X_k \sim N\left(a_k, \sigma_k^2\right), k=1, \cdots, n. c_1, c_2, \cdots, c_n$ 为常数, 记 $T=\sum_{k=1}^n c_k X_k$, 则 $T \sim N\left(\mu, \tau^2\right)$, 其中 $\mu=\sum_{k=1}^n c_k a_k, \tau^2=\sum_{k=1}^n c_k^2 \sigma_k^2$.

证明:利用特征函数证明.

利用多元正态的定义, 可以有

Thr

定理 2. 设随机向量 $X \sim N_p(\mu, \Sigma), A$ 为 $p \times p$ 可逆矩阵, 令 $Y=A X$, 则

$$Y \sim N\left(A \mu, A \Sigma A^{\prime}\right)$$

由正态分布随机变量的变换,可以得到新的分布随机变量:

Def

定义 1. 设 $X_1, X_2, \cdots, X_n$ i.i.d. $\sim N(0,1)$, 则称

$$\xi=\sum_{i=1}^n X_i^2$$

是自由度为 n 的 χ^2 变量, 其分布称为自由度为 n 的 χ^2 分布, 记为 $\xi \sim \chi_n^2$.

ξ 的密度函数 $g_n(x)$ 为

$$g_n(x)= \begin{cases}\frac{1}{2^{n / 2} \Gamma(n / 2)} x^{n / 2-1} e^{-x / 2}, & x>0 \\ 0, & x \leq 0\end{cases}$$

求密度方法: 特征函数法, 归纳法, n 元变换方法.

性质:

(1) 设r.v. $\xi \sim \chi_n^2$, 则 ξ 的c.f. 为 $\varphi(t)=(1-2 i t)^{-\frac{n}{2}}$;
(2) r.v. ξ 的数学期望和方差分别为 $E(\xi)=n, D(\xi)=2 n$.
(3) 设 $Z_1 \sim \chi_{n_1}^2, Z_2 \sim \chi_{n_2}^2$, 且 Z_1 和 Z_2 独立, 则 $Z_1+Z_2 \sim \chi_{n_1+n_2}^2$.

推广: 非中心 χ^2 分布(没整理)

Def

定义 2. 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立, $X_i \sim N(a_i, 1), a_i(i=1,\cdots,n)$ 不全为 0 . 记 $Y=\sum_{i=1}^n X_i^2$\delta=\sqrt{\sum_{i=1}^n a_i^2}

特别当 $\delta=0$ 时称为中心的 χ^2 分布, 即前面所述 χ_n^2 分布.

若 $Y \sim \chi_{n,\delta}^2$,则其密度函数为

$$
\begin{aligned}
g(x) &= \begin{cases} e^{-\delta^2/2} \sum_{i=0}^{\infty} \frac{1}{i!}\left(\frac{\delta^2}{2}\right)^i \frac{x^{i+n/2-1}}{2^{i+n/2}\Gamma(n/2+i)} e^{-x/2}, & x>0 \\ 0, & x \le 0 \end{cases} \\
&= \begin{cases} e^{-\delta^2/2} \sum_{i=0}^{\infty} \frac{(\delta^2/2)^i}{i!} \chi^2(x, 2i+n), & x>0, \\ 0, & x \le 0 \end{cases}
\end{aligned} \tag{1.1}
$$

此处 $\chi^2(x, 2i+n)$ 表示自由度为 $2i+n$ 的 χ^2 密度函数。

非中心 χ^2 密度的计算方法:

令矩阵 A 为正交矩阵, 其第一行元素为 $(a_1/\delta, \cdots, a_n/\delta)$. 从而若 $Y=AX$, 则 $Y_1=\frac{1}{\delta}\sum_{i=1}^n a_i X_i \sim N(\delta, 1)$.

非中心的 χ^2 变量具有下列性质:

(1) 若 $Y \sim \chi_{n,\delta}^2$, 则 Y 的 c.f.为 $\varphi(t)=(1-2it)^{-\frac{n}{2}} e^{\frac{i\delta^2 t}{1-22tt}}$,

(2) 若 $Y \sim \chi_{n,\delta}^2$, 则 $E(Y)=n+\delta^2, \quad D(Y)=2n+\delta^2$,

(3) 若 $Y_1, \cdots, Y_k$ 相互独立, $Y_i \sim \chi_{n_i,\delta_i}^2, i=1,2,\cdots,k$, 则 $\sum_{i=1}^k Y_i \sim \chi_{n,\delta}^2$, 此处 $n=\sum_{i=1}^k n_i$, $\delta=\sqrt{\delta_1^2+\delta_2^2+\cdots+\delta_k^2}$.

(没整理)

由正态分布和 χ^2 分布可以构造一个 **t 分布随机变量**:

Def

定义 3. 设r.v. $X \sim N(0,1), Y \sim \chi_n^2$, 且 X 和 Y 独立, 则称

$$
T=\frac{X}{\sqrt{Y/n}}
$$

为自由为 n 的 t 变量, 其分布称为由为 n 的 t 分布, 记为 $T \sim t_n$.

其密度函数为

$$
t_n(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, -\infty<x<\infty \tag{1.2}
$$

t 变量具有下列的性质:

(1)若r.v. $T \sim t_n$, 则 $E\left(T^r\right)$ 只有当 $r<n(n>1)$ 时存在, 且

$$E\left(T^r\right)=\begin{cases} n^{\frac{r}{2}} \frac{\Gamma\left(\frac{r+1}{2}\right) \Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{1}{2}\right)}, & \text{当 r 为偶数,} \\ 0, & \text{当 r 为奇数.} \end{cases}$$

特别当 $n \geq 2$ 时, $E(T)=0$. 当 $n \geq 3$ 时, $D(T)=\frac{n}{n-2}$.
(2)当 $n=1$ 时 t 分布就是柯西分布, 此时 1.2 变为

$$t_1(x)=\frac{1}{\pi\left(1+x^2\right)},-\infty<x<+\infty$$

(3)当 $n \rightarrow \infty$ 时, t 变量的极限分布为 $N(0,1)$.

非中心 t 分布:(未整理)

定义 4. 设r.v. $X \sim N(\delta, 1), Y \sim \chi_n^2$, 且 X 和 Y 独立, 则称

$$Z=\frac{X}{\sqrt{Y / n}}$$

的分布服从自由度为 n, 非中心参数为 δ 的非中心 t 分布, 记为 $Z \sim t_{n, \delta}$. 特别当 $t=0$ 时的分布称为中心的 t 分布,即前面所述的 t_n 分布.

其密度函数为

$$t_{n, \delta}(x)=\frac{n^{n / 2}}{\sqrt{\pi} \Gamma(n / 2)} \cdot \frac{e^{-\delta^2 / 2}}{\left(n+x^2\right)^{\frac{n+1}{2}}} \sum_{i=0}^{\infty} \Gamma\left(\frac{n+i+1}{2}\right) \frac{(\delta x)^i}{i!}\left(\frac{2}{n+x^2}\right)^{i / 2} \\ -\infty<x<\infty \tag{1.3}$$

密度函数 1.3 的推导方法也是利用求 $r.v.$商的密度函数公式,经过较复杂的计算可求得.

非中心 t 分布的性质如下:

(1) 若 $Z_n \sim t_{n, \delta}$, 则 $Z_n \xrightarrow{\mathscr{L}} N(\delta, 1)$;

(2) 若 $Z_n \sim t_{n, \delta}$, 则有

$$\begin{aligned} & E\left(Z_n\right)=\delta\left(\frac{n}{2}\right)^{\frac{1}{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \quad n \geq 2 \\ & D\left(Z_n\right)=\frac{n\left(1+\delta^2\right)}{n-2}-\frac{\delta^2 n}{2}\left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2, \quad n \geq 3 \end{aligned}$$

χ^2 分布随机变量还可以构造 F 分布随机变量:

(未整理)

Def

定义 5. 设r.v. $X \sim \chi_m^2, Y \sim \chi_n^2$, 且 X 和 Y 独立, 则称

$$F=\frac{X/m}{Y/n}$$

为自由度分别是 m 和 n 的 F 变量, 其分布称为由度分别是 m 和 n 的 F 分布, 记为 $F\sim F_{m,n}$.

其密度函数为

$$f_{m,n}(x)=\begin{cases}\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}m^{\frac{m}{2}}n^{\frac{n}{2}}x^{\frac{m}{2}-1}(n+mx)^{-\frac{m+n}{2}}, & x>0 \quad nbsp;\\ 0, & \text{其它}\end{cases} \tag{1.4}$$

F 变量具有下列的性质:

(1) 若 $Z\sim F_{m,n}$, 则 $1/Z\sim F_{n,m}$.
(2) 若 $Z\sim F_{m,n}$, 则对 $r>0$ 有

$$E\left(X^r\right)=\left(\frac{n}{m}\right)^r\frac{\Gamma\left(\frac{m}{2}+r\right)\Gamma\left(\frac{n}{2}-r\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)},\ \text{当 } 2r<n$$

特别

$$E(X)=\frac{n}{n-2},n>2$$

$$D(X)=\frac{2n^2(m+n-2)}{m(n-2)(n-4)},n>4$$

(3) 若 $T\sim t_n$, 则 $T^2\sim F_{1,n}$
(4) $F_{m,n}(1-\alpha)=1/F_{n,m}(\alpha)$

(未整理)
非中心 F 分布

定义 6. 设r.v. $X\sim\chi^2_{m,\delta}, Y\sim\chi^2_n$, 且 X 和 Y 独立, 则

$$Z=\frac{X/m}{Y/n}$$

的分布称为自由度为 m,n 和非中心参数为 δ 的非中心 F 分布, 记为 $Z\sim F_{m,n;\delta}$. 当 $\delta=0$ 时, 称 Z 的分布为中心的 F 分布, 即前面定义的 $F_{m,n}$.

若 $Z\sim F_{m,n;\delta}$, 则 Z 的密度函数为

$$f_{m,n;\delta}(x)=\begin{cases}\frac{m^{\frac{n}{2}}n^{\frac{m}{2}}}{\Gamma\left(\frac{n}{2}\right)}e^{-\frac{\delta^2}{2}}x^{\frac{m}{2}-1}\sum_{k=0}^{\infty}\frac{\left(\frac{\delta^2mx}{2}\right)^k\Gamma!\left(\frac{m+n}{2}+k\right)}{k!\Gamma\left(\frac{m}{2}+k\right)(mx+n)^{\frac{m+n}{2}+k}}, & x>0 \quad nbsp;\\ 0, & \text{其它}\end{cases} \tag{1.5}$$

非中心 F 分布具有下列性质:

(1) 若 $X\sim t_{n,\delta}$, 则 $X^2\sim F_{1,n;\delta}$.

(2)若 $Z_n\sim F_{m,n;\delta}, n=1,2,\cdots,\delta$ 固定, 则当 $n\to\infty$ 时 $Z_n\xrightarrow{\mathscr{L}}X^2_{m,\delta}$.

(3)若 $Z \sim F_{m,n;\delta}$, 则

$$E(Z) = \frac{n(m+\delta)}{m(n-2)}, \quad 对\ n > 2.$$

$$D(Z) = \frac{2n^2}{m^2(n-2)^2(n-4)}\left[\left(m+\delta^2\right)^2 + (n-2)\left(m+2\delta^2\right)\right], \quad n > 4.$$