三大抽样分布

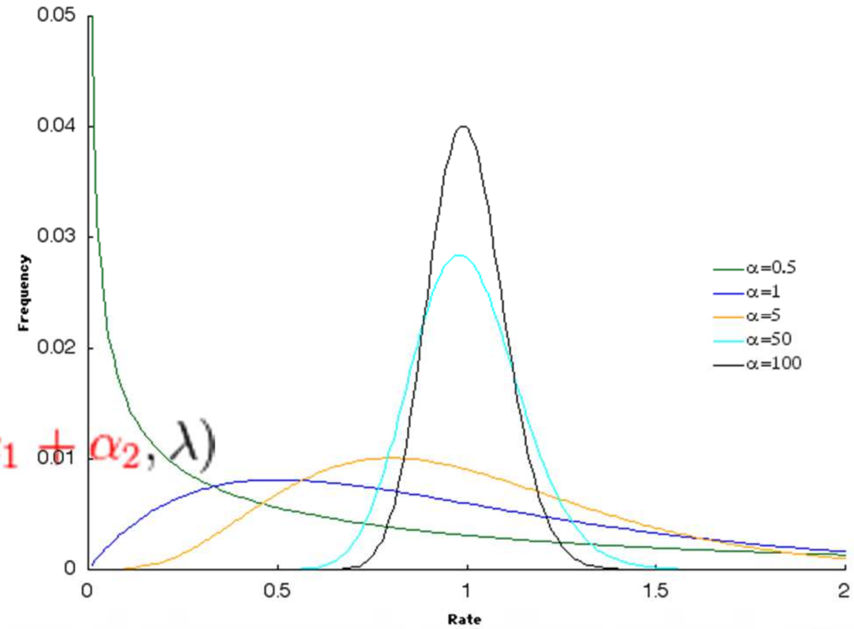
主讲人：刘宏志
liuhz@ss.pku.edu.cn

北京大學

回顾：伽马分布

- 概率密度函数
  - $f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, x \geq 0$
  - 伽马函数：$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} \mathrm{d}u, x > 0$
  - α称为形状参数，λ称为尺度参数
- 性质：
  - 期望：$E(X) = \alpha/\lambda$
  - 方差：$D(X) = \alpha/\lambda^2$
  - 可加性：

$$\begin{cases} r.v. X \sim \Gamma(\alpha_1, \lambda) \\ r.v. Y \sim \Gamma(\alpha_2, \lambda) \end{cases} \Longrightarrow X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$$



- 与指数分布的关系：
  - 当α=1时，伽马分布等价于指数分布

伽马函数

- $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} \mathrm{d}u,\ x>0$
- 性质：广义的阶乘函数
  - $\Gamma(1) = 1$
  - $\Gamma(x+1) = x\Gamma(x)$
  - $\Gamma(n+1) = (n)!$，n为正整数
  - $\Gamma(1/2) = \sqrt{\pi}$

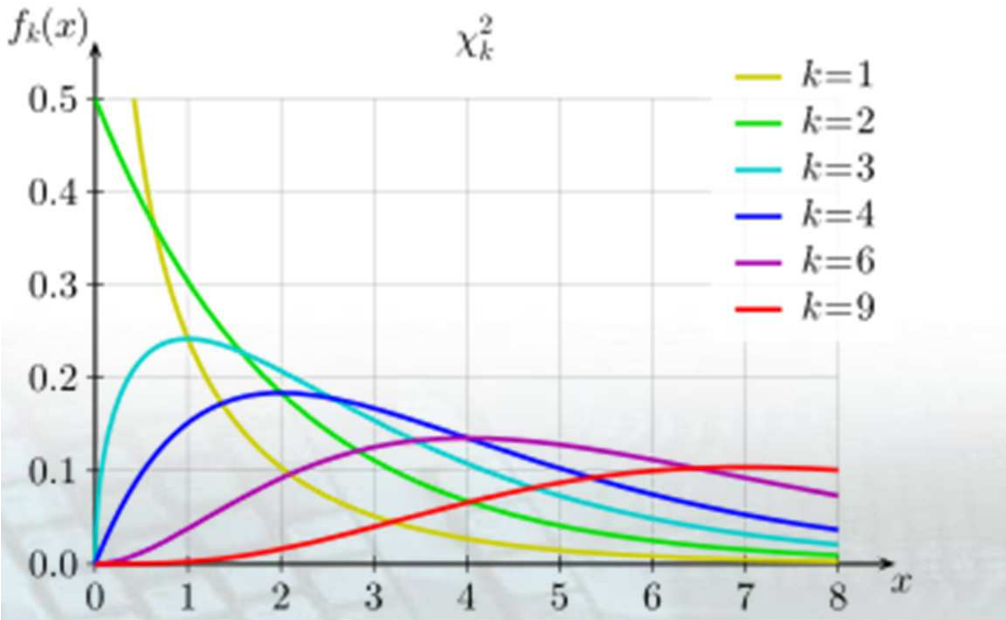
卡方分布

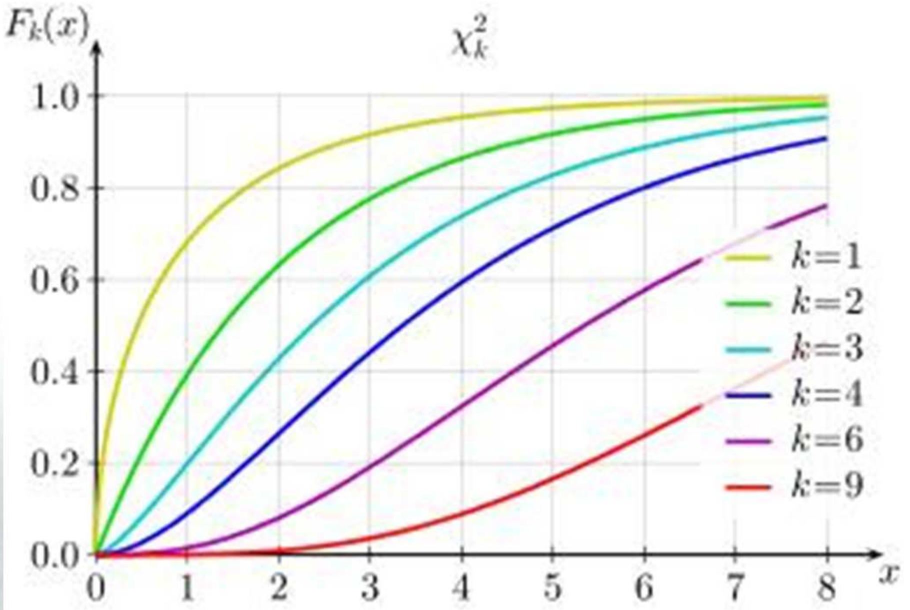
- 若Z是标准正态分布随机变量，即$Z\sim N(0, 1)$，则$U=Z^2$的分布称为自由度为1的卡方分布，记作χ_1^2
- 概率密度函数：$f(u)=\dfrac{u^{-1/2}}{\sqrt{2\pi}}e^{-u/2}$，$u>0$
- $\chi_1^2\sim\mathrm{Ga}\,(1/2, 1/2)$
- 期望：$\mathrm{E}(U)=1$； 方差：$\mathrm{Var}(U)=2$

$$g(x)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x\geq 0$$
$$\Gamma(x)=\int_0^{\infty}u^{x-1}e^{-u}\mathrm{d}u, x>0$$
$$\mathrm{E}(X)=\alpha/\lambda; \quad \mathrm{D}(X)=\alpha/\lambda^2$$

卡方分布

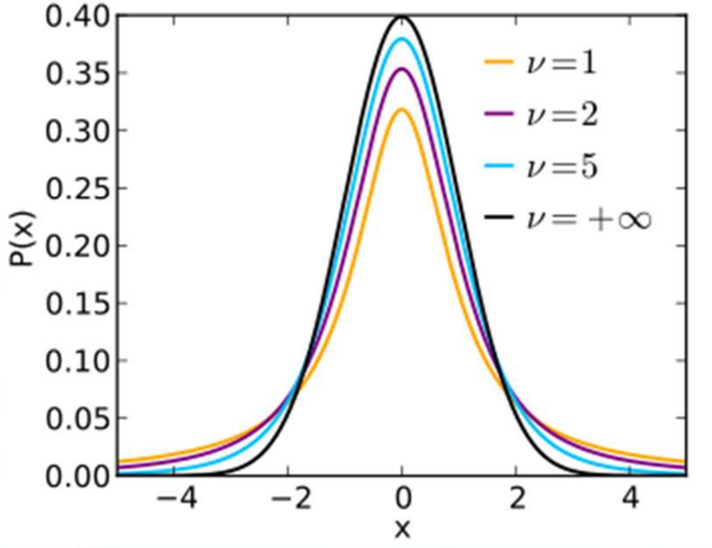
- 若$X_1, X_2, \dots, X_n$是相互独立的且服从标准正态分布的随机变量，则称$\chi_n^2 = X_1^2 + X_2^2 + \dots + X_n^2$所服从的分布为自由度为$n$的卡方分布
- 若$U_1, U_2, \dots, U_n$是相互独立的自由度为1的卡方随机变量，则称$V=U_1+U_2+\dots+U_n$所服从的分布为自由度为n的卡方分布，记作χ_n^2
- 概率密度函数：$f(v) = \dfrac{v^{(n/2)-1}}{2^{n/2}\Gamma(n/2)} e^{-v/2}$，$v>0$
- 性质：$\chi_n^2 \sim Ga(n/2, 1/2)$
  - 期望：$E(V)=n$; 方差：$Var(V)=2n$
  - 可加性：$U \sim \chi^2_n, V \sim \chi^2_m$，且$U$和$V$独立，则$U+V \sim \chi^2_{m+n}$





t分布

- 若$Z \sim N(0,1)$，$U \sim \chi_n^2$，且Z和U独立，则称随机变量$T=Z/\sqrt{U/n}$服从的分布是自由度为n的t分布
- 密度函数：$f(t)=\dfrac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)}(1+\dfrac{t^2}{n})^{-(n+1)/2}$
- 性质：
  - 对称性：t分布关于零点对称，$f(t)=f(-t)$, $E(T)=0$
  - 当自由度趋向无穷时($n>30$)，t分布趋向于标准正态分布

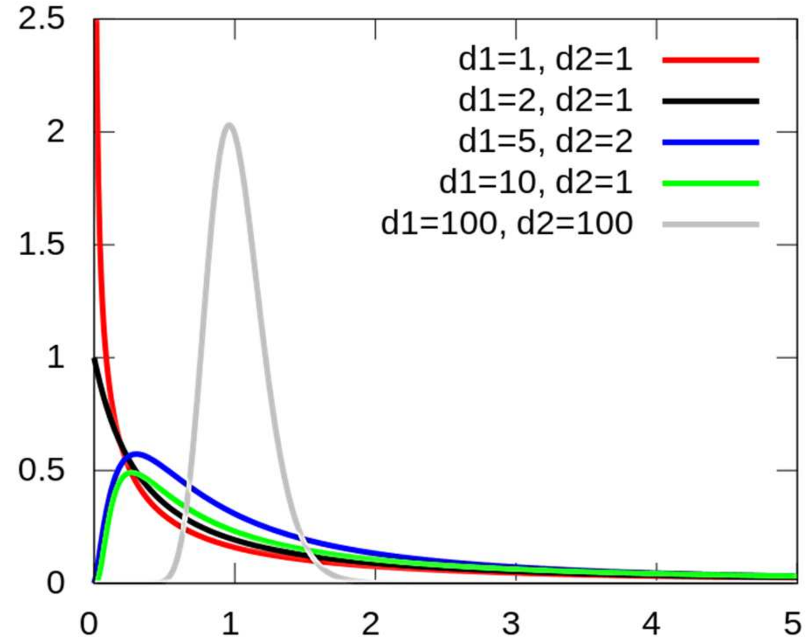


F分布

- 若$U \sim \chi_m^2$， $V \sim \chi_n^2$，且U和V相互独立，则称$W=\frac{U/m}{V/n}$服从的分布是自由度为m和n的F分布，记作$F_{m,n}$
- 密度函数：

$$f(w)=\frac{\Gamma\left[\frac{(m+n)}{2}\right]}{\Gamma\left(\frac{m}{2\Gamma\left(\frac{n}{2}\right)}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}}w^{\frac{m}{2}-1}\left(1+\frac{m}{n}w\right)^{-(m+n)/2}$$

- 性质：
  - 当$n>2$时，$E(W)=n/(n-2)$
  - 若$W \sim F_{m,n}$，则$W^{-1} \sim F_{n,m}$



回顾：样本均值和样本方差

- 设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$的样本,
- 样本均值： $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ ；
- 样本方差：$S^2 = \frac{1}{n-1}\sum_{i=}^{n}(X_i - \bar{X})^2$
- $\bar{X} \sim N(\mu, \ \sigma^2/n)$，即 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- $\bar{X}$和S^2是相互独立的

- 定理1: 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$的样本,则

$$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi_n^2$$

- 定理2:设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$的样本, S^2为样本方差，则

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

北京大学

推论1 设 $X_1, X_2, \cdots, X_n$ 是总体 $N(\mu, \sigma^2)$ 的样本, $\overline{X}, S^2$ 分别是样本均值和样本方差, 则有

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1).$$

证明 因为 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$,

且两者独立, 由 t 分布的定义知

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \Bigg/ \sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}} \sim t(n-1).$$

设 $X_1, X_2, \cdots, X_n$ 是总体 $N(\mu,\sigma^2)$ 的样本,

$\overline{X}, S^2$ 分别是样本均值和样本方差, 则有

结论

(1) $\overline{X} \sim N(\mu, \sigma^2 / n)$. 即 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

(2) $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

(3) $\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$

(4) $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2} \sim \chi^2(n-1)$

(5) $\overline{X}$ 与 S^2 相互独立