Statistics

Sample

$X_1,\dots,X_n \overset{i.i.d}{\sim} (\mu,\sigma^2)$

1. **sample mean**

   $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \bar{X} \sim (\mu, \frac{\sigma^2}{n})$

   $E(X^2) = E(X)^2 + Var(X) = \mu^2 + \sigma^2$

   $E(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$

2. **sample variance**

   $S^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}, E(S^2)$

   $\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$

   $\sum_{i=1}^{n} X_i^2 = n\bar{X}^2 + (n-1)S^2$

   $\bar{X}$ and S^2 are independent when Normal distribution

Parameter Estimation

$X \sim \begin{cases} f(x|\theta) & continues \\ P(x|\theta) & discrete \end{cases}$, depending on $\theta \in \Theta$

Estimator

1. an **estimator** $\hat{\theta}$:

- is function of samples($X_1,\dots,X_n$)
- has distribution depending on θ

  $X \sim (\mu,\sigma^2), E(\bar{X}) = \mu, E(S^2) = \sigma^2$

2. **unbiased**: if $E(\hat{\theta}) = \theta$
3. **MSE**(Mean Square Error)

   $MSE = E((\hat{\theta}-\theta)^2) = Var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2 = Var(\hat{\theta}) + (bias(\hat{\theta}))^2$

   if $\hat{\theta}$ is unbiased, $MSE = Var(\hat{\theta})$

4. **SE**(Standard Error)
   1. 根据$\hat{\theta}$的分布或者近似分布，求$Var(\hat{\theta}) = V(\theta)$
   2. 用$\hat{\theta}$替换方差$V(\theta)$中的θ，$SE = \sqrt{V(\hat{\theta})}$

MLE(Maximum Likelihood Estimation)

- 求MLE
  1. 写出Likelihood function $L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$
  2. 求$logL(\theta) = \sum_{i=1}^{n} logf(x_i|\theta)$
  3. 求一阶导: $\frac{\partial logL(\theta)}{\partial\theta} = 0 \Rightarrow \theta = \theta(X)$
  4. 求二阶导：$\frac{\partial^2 logL(\theta)}{\partial\theta^2}|_{\theta=\theta(X)} < 0$
  5. 得到MLE：$\hat{\theta} = \theta(X)$
- **Fisher Information**

  $I(\theta) = -E((\frac{\partial^2}{\partial\theta^2}logf(x|\theta))^2) = E((\frac{\partial}{\partial\theta}logf(x|\theta))^2)$

此处X是一个样本的分布

- **Asymptotic Distribution**

  $\hat{\theta} \sim N(\theta_0, \frac{1}{nI(\theta_0)})$

  $SE = \frac{1}{\sqrt{nI(\hat{\theta})}}$

  $\alpha - CI : \hat{\theta} \pm Z_{\frac{\alpha}{2}} \frac{1}{\sqrt{nI(\hat{\theta})}}$

Moments Estimation

1. 求期望$\mu = E(X) = h(\theta)$
2. 用$\bar{X}$代替μ，通过$\bar{X} = E(X) = h(\theta)$解得$\hat{\theta} = h^{-1}(\bar{X})$

Confidence Interval

1. 找到一个统计量$T = T(X, \theta) \sim F$
2. 给定α，得到T的$CI = [F^{-1}_{\frac{\alpha}{2}}，F^{-1}_{1-\frac{\alpha}{2}}]$
3. 通过$F^{-1}_{\frac{\alpha}{2}} \leq T(\theta) \leq F^{-1}_{1-\frac{\alpha}{2}}$解得$\theta$的$CI$

条件	参数	Statistics	CI
unknown μ **known** σ^2	μ	$T = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$	$\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
unknown μ **unknown** σ^2	μ	$T = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t_{n-1}$	$\bar{X} \pm t_{n-1}(\frac{\alpha}{2}) \frac{S}{\sqrt{n}}$

Central limit theorem

1. 定理
   - 独立同分布

     $X_i \overset{i.i.d}{\sim} (\mu, \sigma^2)$

     $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

     $S_n \sim N(n\mu, n\sigma^2)$

   - 独立不同分布

     $X_i \sim (\mu_i, \sigma_i^2)$ are independent

     $S_n \sim N(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2)$

     $\bar{X} \sim N(\frac{\sum_{i=1}^{n} \mu_i}{n}, \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2})$

2. 求概率/置信区间

   对于$S_n = \sum_{i=1}^{n} X_i$，利用$\frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \sim N(0,1)$($\bar{X}$同理)

   - 求概率

     $P(S_n \leq x) = P(\frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \leq \frac{x - E(S_n)}{\sqrt{Var(S_n)}}) = \Phi(\frac{x - E(S_n)}{\sqrt{Var(S_n)}})$

   - 求置信区间

     $\alpha - CI = E(S_n) \pm Z_{\frac{\alpha}{2}} \sqrt{Var(S_n)}$

Rejection Sampling

1. 已知可从分布$X \sim H$产生随机数，构造从$X \sim F$中产生随机数

   找到常数$c : f(x) \leq ch(x), \forall x$

   - 从H中产生一个随机数X
   - 以$\frac{f(X)}{ch(X)}$概率保留X

   对于离散分布，用PMF代替PDF

Efficiency为$\frac{1}{c}$，c越小效率越高

2. 任意随机变量X，其CDF：F满足

   $F(X) \sim U(0,1)$

   $P(F(X) \le x) = P(X \le F^{-1}(x)) = F(F^{-1}(x)) = x \Rightarrow f(F(X)) = 1, F(X) \in [0,1]$

Hypothesis Testing

Concepts

- **Null Hypothesis**: H_0 vs **Alternative Hypothesis** : H_1
- **Test Statistic**: T
  - 是随机变量，原假设下分布 $T|H_0 \sim F$
  - 给定α，可根据原假设求RR

    $$RR = \begin{cases} \{|T| > F_{1-\frac{\alpha}{2}}\} & H_0 : \theta = \theta_0 \\ \{T > F_{1-\alpha}\} & H_0 : \theta \le \theta_0 \\ \{T < -F_{\alpha}\} & H_0 : \theta \ge \theta_0 \end{cases}$$
  - 是样本的函数$T(X)$
  - reject H_0 if $T(X) \in RR$
- **Rejection Region**: RR
- **Type I Error** (**Significance Level**): $\alpha = P(T \in RR|H_0)$

  rejecting H_0 when H_0 is true
- **Type II Error**: $\beta = P(T \notin RR|H_1)$

  accepting H_0 when H_0 is false

  for **specific value** of the alternative hypothesis: $\beta(\theta_1) = P(T \notin RR|\theta = \theta_1 \in \Theta_1)$
- **Statistical Power**: $1 - \beta$
- **p-value**: $P(T \text{ is extreme than } T_{obs}|H_0)$
  1. 找分布和统计量的样本观测值

     $T|\theta = \theta_0 \sim F, T_0 = T(X)$
  2. 求$p\text{-}value$

     $$p\text{-}value = \begin{cases} P(|T| > T_0|T \sim F) & H_0 : \theta = \theta_0 \\ P(T > T_0|T \sim F) & H_0 : \theta \le \theta_0 \\ P(T < T_0|T \sim F) & H_0 : \theta \ge \theta_0 \end{cases}$$
  3. 给定 α

     $$\begin{cases} \text{reject } H_0 & p\text{-}value \le \alpha \\ \text{accept } H_0 & p\text{-}value > \alpha \end{cases}$$

Likelihood Ratio Tests(Simple Hypothesis)

$H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$

1. 写出 Likelihood function: $L(\theta) = \prod\limits_{i=1}^{n} f(x_i|\theta)$
2. 求 Likelihood ratio: $\Lambda(X) = \frac{L(\theta_0)}{L(\theta_1)}$
3. threshold C，得到拒绝域$RR = \{\Lambda(X) \le C\}$

   若C未知，给定α，令$\alpha = P(\Lambda(X) \le C|X \sim F_{\theta_0}) \overset{\text{通过}\Lambda\text{反解出}X}{=} P(X \le \Lambda^{-1}(C)|X \sim F_{\theta_0})$可求得C
4. 做决策，$\begin{cases} \text{reject } H_0 & \Lambda \le C \\ \text{accept } H_0 & \Lambda > C \end{cases}$

Generalized Likelihood Ratio Tests

$H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$

$\Theta = \Theta_0 \cup \Theta_1, d = dim(\Theta), d_0 = dim(\Theta_0)$

1. 写出 Likelihood function: $L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$
2. 求Θ_0和Θ下的MLE:

   $\hat{\theta_0} = \underset{\theta\in\Theta_0}{argMax} L(\theta)$

   $\hat{\theta} = \underset{\theta\in\Theta}{argMax} L(\theta)$
3. 求 Likelihood ratio: $\Lambda = \frac{L(\hat{\theta_0})}{L(\hat{\theta})}$
4. 求分布：$-2log\Lambda \sim \mathcal{X}^2_{d-d_0}$
5. 给定α
   - $RR = \{-2log\Lambda > F_{\mathcal{X}^2_{d-d_0}}(1-\alpha)\}, \begin{cases} \text{reject } H_0 & -2log\Lambda > F_{\mathcal{X}^2_{d-d_0}}(1-\alpha) \\ \text{accept } H_0 & -2log\Lambda \le F_{\mathcal{X}^2_{d-d_0}}(1-\alpha) \end{cases}$
   - $p\text{-}value = 1 - F^{-1}_{\mathcal{X}^2_{d-d_0}}(-2log\Lambda), \begin{cases} \text{reject } H_0 & p\text{-}value \le \alpha \\ \text{accept } H_0 & p\text{-}value > \alpha \end{cases}$

Pearson's $\mathcal{X}^2$ Test

计算 $\mathcal{X}^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij}-E_{ij})^2}{E_{ij}} \sim \mathcal{X}^2_{d-d_0}$:

O_{ij}：第i行第j列观测值，$E_{ij} = n \cdot \hat{p_{ij}}$：第$i$行第$j$列的期望

$d = dim(\Theta), d_0 = dim(\Theta_0)$

$RR = \{\mathcal{X}^2 > F_{\mathcal{X}^2_{d-d_0}}(1-\alpha)\}, \begin{cases} \text{reject } H_0 & \mathcal{X}^2 > F_{\mathcal{X}^2_{d-d_0}}(1-\alpha) \\ \text{accept } H_0 & \mathcal{X}^2 \le F_{\mathcal{X}^2_{d-d_0}}(1-\alpha) \end{cases}$

$p\text{-}value = 1 - F^{-1}_{\mathcal{X}^2_{d-d_0}}(\mathcal{X}^2), \begin{cases} \text{reject } H_0 & p\text{-}value \le \alpha \\ \text{accept } H_0 & p\text{-}value > \alpha \end{cases}$

- 2×2 Case

	1	**2**	
1	n_{11}, p_{11}	n_{12}, p_{12}	R_1
2	n_{21}, p_{21}	n_{22}, p_{22}	R_2
	C_1	C_2	n

  1. 根据H_0下的条件，设参数$p, \ldots,$ 简化$p_{ij}, L(p) = \binom{n}{n_{11}, n_{12}, n_{21}, n_{22}} p_{11}^{n_{11}} p_{12}^{n_{12}} p_{21}^{n_{21}} p_{22}^{n_{22}}$
  2. 求MLE并计算每个格子的概率估计: $\hat{p}, \ldots \to \hat{p_{11}}, \hat{p_{12}}, \hat{p_{21}}, \hat{p_{22}}$
- Independent Test

  $H_0: p_{ij} = p_i \cdot q_j$, 检验行和列是否独立

	1	**...**	j	**...**	c	$\sum$
1	n_{11}	...	n_{1j}	...	n_{11}	R_1, p_1
...	...		...		...	...
i	n_{i1}	...	n_{ij}	...	n_{ic}	R_i, p_i
...	...		...		...	...
r	n_{r1}	...	n_{rj}	...	n_{rc}	R_r, p_r
$\sum$	C_1, q_1	...	C_j, q_j	...	C_c, q_c	n

  $\hat{p_i} = \frac{R_i}{n}, \hat{q_j} = \frac{C_j}{n}, E_{ij} = \frac{R_i \times C_j}{n}, df = (r-1)(c-1)$
- Consistent Test

  $H_0: (p_{i1}, \ldots, p_{ic}) = (p_1, \ldots, p_c), i = 1, \ldots, r$ 检验列在不同的行分类下是否有一致分布

	1	...	j	...	c	$\sum$
1	n_{11}	...	n_{1j}	...	n_{11}	R_1
...	...		...		...	...
i	n_{i1}	...	n_{ij}	...	n_{ic}	R_i
...	...		...		...	...
r	n_{r1}	...	n_{rj}	...	n_{rc}	R_r
$\sum$	C_1, p_1	...	C_j, p_j	...	C_c, p_c	n

- known $(p_1, \dots, p_c)$

  $E_{ij} = R_i \times p_j, df = r(c-1)$
- unknown $(p_1, \dots, p_c)$

  $\hat{p_i} = \frac{C_i}{n}, E_{ij} = \frac{R_i \times C_j}{n}, df = (r-1)(c-1)$

Normal Distribution

$X_i \sim N(\mu, \sigma^2)$

参数情况	H_0	H_1	Statistics	RR	p-$value$
unknown μ **known** σ^2	$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ $T\|_{\mu=\mu_0} \sim N(0,1)$	$\|T\| > Z_{\frac{\alpha}{2}}$	$P(\|T\| > \frac{\sqrt{n}\|\bar{X}-\mu_0\|}{\sigma}\|H_0)$
	$\mu \leq \mu_0$	$\mu > \mu_0$		$T > Z_\alpha$	$P(T > \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}\|H_0)$
	$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -Z_\alpha$	$P(T < \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}\|H_0)$
unknown μ **unknown** σ^2	$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \frac{\sqrt{n}(\bar{X}-\mu)}{S}$ $T\|_{\mu-\mu_0} \sim t_{n-1}$	$\|T\| > t_{n-1}(\frac{\alpha}{2})$	$P(\|T\| > \frac{\sqrt{n}\|\bar{X}-\mu_0\|}{S}\|H_0)$
	$\mu \leq \mu_0$	$\mu > \mu_0$		$T > t_{n-1}(\alpha)$	$P(T > \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}\|H_0)$
	$\mu \geq \mu_0$	$\mu < \mu_0$		$T < -t_{n-1}(\alpha)$	$P(T < \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}\|H_0)$

Bayesian Inference

Hierarchical Model

- 参数$\theta \in \Theta$看作一个随机变量，有 **prior distribution**: $\theta \sim f_\Theta(\theta)$

  如果为**discrete**，用PMF代替PDF
- 样本X基于θ的**conditional distribution**为$X_1, \dots, X_n | \theta \sim f_{x|\theta}(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} f_{x|\theta}(x_i|\theta)$

Related Distribution

- 样本X和参数θ的**joint distribution**为$f(\theta, x_1, \dots, x_n) = f_\theta(\theta) f_{x|\theta}(x_1, \dots, x_n|\theta) = f_\theta(\theta) \prod_{i=1}^{n} f_{x|\theta}(x_i|\theta)$
- 样本X的**marginal distribution**(对θ求积分)为

  $f_x(x_1, \dots, x_n) = \int_\Theta f(\theta, x_1, \dots, x_n) d\theta = \int_\Theta f_{x|\theta}(x_1, \dots, x_n|\theta) f_\theta(\theta) d\theta$
- 参数θ基于样本X的**posterior distribution**为$f_{\theta|x}(\theta|x_1, \dots, x_n) = \frac{f(\theta, x_1, \dots, x_n)}{f_x(x_1, \dots, x_n)} = \frac{f_{x|\theta}(x_1, \dots, x_n|\theta) f_\theta(\theta)}{f_x(x_1, \dots, x_n)} = \frac{f_{x|\theta}(x_1, \dots, x_n|\theta) f_\theta(\theta)}{\int_\Theta f(\theta, x_1, \dots, x_n) d\theta}$

  求后验分布如果很复杂，如$Beta, Gamma$等分布，用$kernel$：

  $f_{\theta|x}(\theta|x_1, \dots, x_n) \propto f_{x|\theta}(x_1, \dots, x_n|\theta) f_\theta(\theta)$

  只保留含有θ的部分($kernel$)，然后根据θ的表达式找对应的分布类型以及参数

Credible Interval

1. 求**posterior distribution** $f_{\theta|x}(\theta|x_1, \dots, x_n)$ with CDF $F_{\theta|x}$

2. 给定α，区间为$[F^{-1}_{\theta|x}(\frac{\alpha}{2}), F^{-1}_{\theta|x}(1-\frac{\alpha}{2})]$

   若后验分布如果很复杂，如$Beta, Gamma$等分布，而题目要求求出具体数，可用正态分布近似，其均值和方差为后验分布的均值和方差

Parameter Estimation

先求**posterior distribution** $f_{\theta|x}(\theta|x_1,\dots,x_n)$，然后基于此分布估计

1. **posterior mean**：后验分布求期望，得$\hat{\theta_E}=E(\theta|X_1,\dots,X_n)$
2. **posterior mode**：后验分布求MLE，得$\hat{\theta}$

New Observation

观测新的样本$X_{n+1}|\theta\sim F(\theta)$，求$X_{n+1}|X_1,\dots,X_n$

用tower law以及$X_{n+1}|(X_1,\dots,X_n,\theta)=X_{n+1}|\theta$，即在固定$\theta$下，$X_{n+1}$与$X_1,\dots X_n$无关

$P(X_{n+1}|X_1,\dots,X_n)=E(P(X_{n+1}|\theta)|X_1,\dots,X_n)$

$E(h(X_{n+1})|X_1,\dots,X_n)=E(E(h(X_{n+1})|\theta)|X_1,\dots,X_n)$

Random Variable

Discrete Distribution

Distribution	Poisson: $X\sim Pois(\lambda)$	Bernoulli: $X\sim Ber(p)$	Binomial Distribution: $X\sim B(n,p)$	Geometric Distribution: $X\sim Geo(p)$	Negative binomial: $X\sim NB(k,p)$	Multinomial: $X=(X_1,\dots,X_k)\sim M(n,p_1,\dots,p_k)$
Parameters	$\lambda>0$	$0\le p\le 1$	$n\in N, 0\le p\le 1$	$p\in(0,1)$	$k>0, p\in(0,1)$	$n>0,\sum_{i=1}^{k}p_i=1$
Support	$x\in N^+$	$x\in\{0,1\}$	$X\in\{0,1,\dots,n\}$	$X\in\{1,2,\dots\}$	$X\in\{k,k+1,\dots\}$	$X_i\in\{0,\dots,n\}, i=1,\dots,k$ $\sum_{i=1}^{k}X_i=n$
PMF	$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,\dots$	$P(X=k)=p^k(1-p)^{1-k}$ $k=0,1$	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ $0\le k\le n$	$P(X=k)=p(1-p)^{k-1}$ $k\ge 1$	$P(X=n)=\binom{n-1}{k-1}p^k(1-p)^{n-k}$ $n\ge k$	$P(X_1=n_1,\dots,X_k=n_k)=$ $\binom{n}{n_1,\dots,n_k}\prod_{i=1}^{k}p_i^{n_i}$
CDF	-	-	-	$P(X<=k)=1-(1-p)^k$	-	-
Mean	$E(X)=\lambda$	$E(X)=p$	$E(X)=np$	$E(X)=\frac{1}{p}$	$E(X)=\frac{k}{p}$	$E(X_i)=np_i$
Variance	$Var(X)=\lambda$	$Var(X)=p(1-p)$	$Var(X)=np(1-p)$	$Var(X)=\frac{1-p}{p^2}$	$Var(X)=\frac{k(1-p)}{p^2}$	$Var(X_i)=np_i(1-p_i)$
MLE	$\hat{\lambda}=\bar{X}$	$\hat{p}=\bar{X}$	-	$\hat{p}=\frac{1}{\bar{X}}$	$\hat{p}=\frac{k}{\bar{X}}$	$\hat{p}=\frac{X_i}{n}$
Fisher Info (Single Sample)	$I(\lambda)=\frac{1}{\lambda}$	$I(p)=\frac{1}{p(1-p)}$	$I(p)=\frac{n}{p(1-p)}$	$I(p)=\frac{1}{p^2(1-p)}$	$I(p)=\frac{k}{p^2(1-p)}$	-
Others			$X=\sum_{i=1}^{n}Y_i, Y_i\overset{i.i.d}{\sim}Ber(p)$	Memoryless: $P(X>m+n\mid X>n)=P(X>m)$	$NB(1,p)=Geo(p)$	$\binom{n}{n_1,\dots,n_k}=\frac{n!}{\prod_{i=1}^{k}n_i!}$

Continues Distribution

Distribution	Uniform: $X\sim U(a,b)$	Exponential: $X\sim Exp(\lambda)$	Normal Distribution: $X\sim N(\mu,\sigma^2)$
Parameters	$-\infty<a<b<\infty$	$\lambda>0$	$\mu\in R,\sigma^2>0$
Support	$x\in[a,b]$	$x\ge 0$	$x\in R$
PDF	$f(x)=\begin{cases}\frac{1}{b-a} & x\in[a,b]\\ 0 & otherwise\end{cases}$	$f(x)=\lambda e^{-\lambda x}$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
CDF	$f(x)=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & x\in[a,b]\\ 1 & x>b\end{cases}$	$F(x)=1-e^{-\lambda x}$	-
Mean	$E(X)=\frac{a+b}{2}$	$E(X)=\frac{1}{\lambda}$	$E(X)=\mu$
Variance	$Var(X)=\frac{(b-a)^2}{12}$	$Var(X)=\frac{1}{\lambda^2}$	$Var(X)=\sigma^2$
MLE	$\begin{cases}\hat{b}=X_{(n)} & \text{known } a, \text{unknown } b\\ \hat{a}=X_{(1)} & \text{unknown } a, \text{known } b\end{cases}$	$\hat{\lambda}=\frac{1}{\bar{X}}$	$logL(\mu,\sigma^2)=-\frac{n}{2}ln(2\pi)-\frac{n}{2}ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$ $\begin{cases}\hat{\mu}=\bar{X} & \text{unknown } \mu, \text{known } \sigma^2\\ \hat{\sigma}^2=\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n} & \text{known } \mu, \text{unknown } \sigma^2\\ \hat{\mu}=\bar{X},\hat{\sigma}^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n} & \text{unknown } \mu, \text{unknown } \sigma^2\end{cases}$
Fisher Info (Single Sample)	-	$I(\lambda)=\frac{1}{\lambda^2}$	$\begin{cases}I(\mu)=\frac{1}{\sigma^2} & \text{unknown } \mu, \text{known } \sigma^2\\ I(\sigma^2)=\frac{1}{2\sigma^4} & \text{known } \mu, \text{unknown } \sigma^2\end{cases}$
Others	$U\sim U(0,1), X\sim U(a,b), X=a+(b-a)U$	- Memoryless: $P(X>s+t\mid X>t)=P(X>s)$ - Order Statistics: $X_{(1)}\sim Exp(n\lambda)$ - Scaling: $aX\sim Exp(\lambda/a), a>0$ - $Exp(\frac{1}{2})=\mathcal{X}^2_2$	- $X\sim N(\mu,\sigma^2), aX+b\sim N(a\mu+b,a^2\sigma^2)$ - $X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$, $aX+bY\sim N(a\mu_1+b\mu_2,a^2\sigma_1^2+b^2\sigma_2^2+2ab\sigma_1\sigma_2\rho)$ $1-\Phi(A)=\Phi(-A)$

Distribution	Gamma: $X\sim\Gamma(\alpha,\ \beta)$	Beta: $X\sim Beta(\alpha,\beta)$	Chi-squared: $X\sim\mathcal{X}^2_n$	Student's *t*: $X\sim t_n$
Parameters	$\alpha>0,\beta>0$	$\alpha>0,\beta>0$	$n\in N^+$	$n>0$
Support	$X>0$	$X\in[0,1]$	$x\in[0,\infty)$	$x\in(-\infty,\infty)$
PDF	$f(x)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	$f(x)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$	$f(x)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{\frac{x}{2}}$	-
CDF	-	-	-	-
Mean	$E(X)=\frac{\alpha}{\beta}$	$E(X)=\frac{\alpha}{\alpha+\beta}$	$E(X)=n$	$E(X)=0$
Variance	$Var(X)=\frac{\alpha}{\beta^2}$	$Var(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$Var(X)=2n$	$Var(X)=\frac{n}{n-2}, n>2$
MLE	-	-	-	-
Fisher Info (Single Sample)	-	-	-	-
Others	- $X_i\overset{indep}{\sim}\Gamma(\alpha_i,\beta)$, $\sum X_i\sim\Gamma(\sum\alpha_i,\beta)$ - $\Gamma(1,\lambda)=Exp(\lambda)$ - $\Gamma(\frac{n}{2},\frac{1}{2})=\mathcal{X}^2_n$		- $Z_i\overset{i.i.d}{\sim}N(0,1)$, $V=\sum Z_i\sim\mathcal{X}^2_n$ - $\mathcal{X}^2_n=\Gamma(\frac{n}{2},\frac{1}{2})$	$Z\sim N(0,1)$ independent with $V\ \sim\mathcal{X}^2_n$, $T=\frac{Z}{\sqrt{\frac{V}{n}}}\sim t_n$