



信阳师范学院  
数学与统计学院  
SCHOOL OF MATHEMATICS AND STATISTICS

# 第10章 数据统计分析

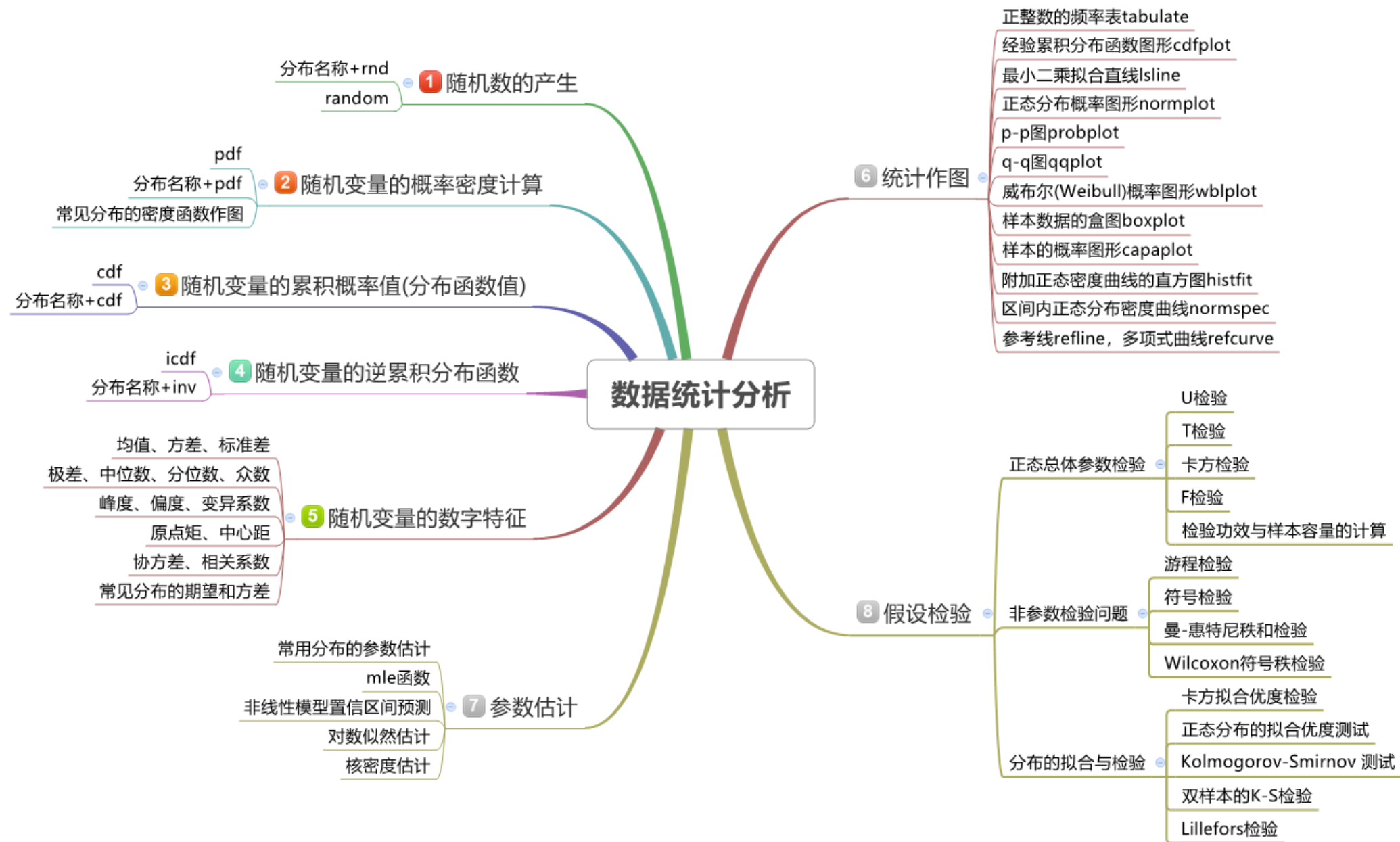


讲授人：牛言涛



日期：2020年4月6日

# 第10章 数据统计分析知识点思维导图



# 1. 常见分布随机数产生表

函数名	调用形式	注 释
unifrnd	unifrnd ( A,B,m,n)	[A,B]上均匀分布(连续) 随机数
unidrnd	unidrnd(N,m,n)	均匀分布（离散）随机数
exprnd	exprnd(Lambda,m,n)	参数为Lambda的指数分布随机数
normrnd	normrnd(MU,SIGMA,m,n)	参数为MU， SIGMA的正态分布随机数
chi2rnd	chi2rnd(N,m,n)	自由度为N的卡方分布随机数
trnd	trnd(N,m,n)	自由度为N的t分布随机数
frnd	frnd(N <sub>1</sub> , N <sub>2</sub> ,m,n)	第一自由度为N <sub>1</sub> ,第二自由度为N <sub>2</sub> 的F分布随机数
gamrnd	gamrnd(A, B,m,n)	参数为A, B的 分布 $\gamma$ 随机数
betarnd	betarnd(A, B,m,n)	参数为A, B的 分布 $\beta$ 随机数
lognrnd	lognrnd(MU, SIGMA,m,n)	参数为MU, SIGMA的对数正态分布随机数

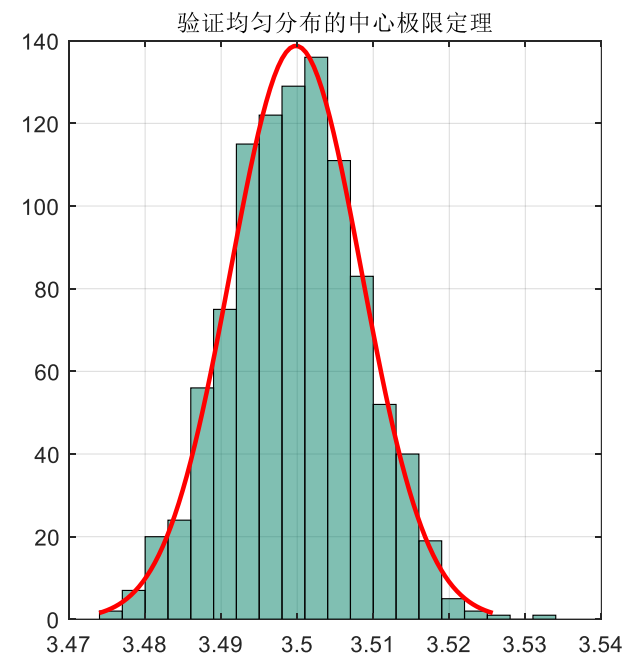
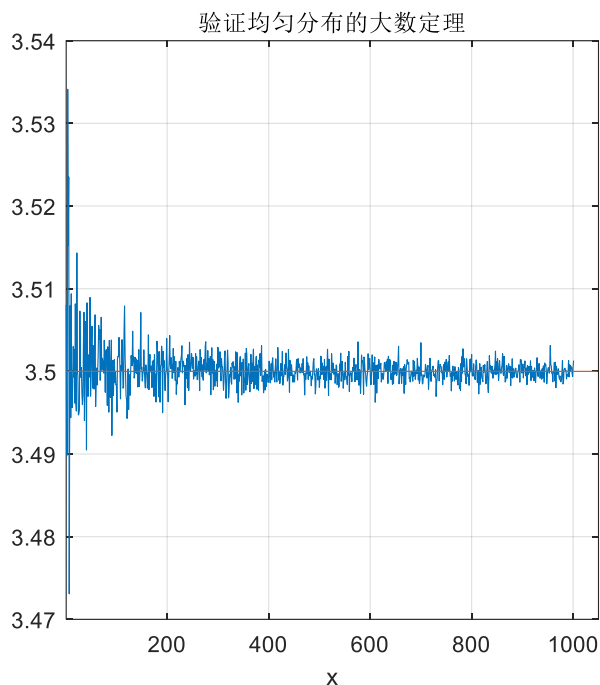
# 1. 常见分布随机数产生表

函数名	调用形式	注 释
nbirrnd	nbirrnd(R, P,m,n)	参数为R, P的负二项式分布随机数
ncfrnd	ncfrnd(N <sub>1</sub> , N <sub>2</sub> , delta,m,n)	参数为N <sub>1</sub> , N <sub>2</sub> , delta的非中心F分布随机数
nctrnd	nctrnd(N, delta,m,n)	参数为N, delta的非中心t分布随机数
ncx2rnd	ncx2rnd(N, delta,m,n)	参数为N, delta的非中心卡方分布随机数
raylrnd	raylrnd(B,m,n)	参数为B的瑞利分布随机数
weibrnd	weibrnd(A, B,m,n)	参数为A, B的韦伯分布随机数
binornd	binornd(N,P,m,n)	参数为N, p的二项分布随机数
geornd	geornd(P,m,n)	参数为 p的几何分布随机数
hygernd	hygernd(M,K,N,m,n)	参数为 M, K, N的超几何分布随机数
poissrnd	poissrnd(Lambda,m,n)	参数为Lambda的泊松分布随机数

## 2. 均匀分布随机数验证大数定理和中心极限定理

```
meanval = mean(unifrnd(2,5,1,500));  
for k = 1000:1000:1000000 %大数定理验证  
    R = unifrnd(2,5,1,k);  
    meanval = [meanval,mean(R)];  
end  
for i = 1:1000 %中心极限定理验证  
    R = unifrnd(2,5,1,10000);  
    clt(i) = mean(R); %Central Limit Theorem  
end  
subplot(1,2,1); plot(meanval)  
hold on; grid on  
ezplot('3.5',[1,1050])  
subplot(1,2,2); histfit(clt,20)  
colormap(summer); alpha(0.5)  
grid on
```

- 简单而言，大数定律讲的是样本均值收敛到总体均值（期望）。
- 中心极限定理，当样本量足够大时，样本均值的分布服从正态分布。



### 3. 二项分布与泊松分布



```
X = binornd(100,0.5,5000,1);
```

```
X_mean = mean(X) %求二项分布的期望np
```

```
X_var = var(X) %方差npq
```

```
subplot(1,2,1); boxplot(X,1) %绘制盒图
```

```
grid on; title('二项分布的盒图')
```

```
subplot(1,2,2)
```

```
k = binornd(1000,0.005,5000,1); k = unique(k);
```

```
% n =1000, p = 0.005, 每次实验事件A发生的次数为k
```

```
pb = binopdf(k,1000,0.005);
```

```
plot(k,pb,'b.','MarkerSize',10)
```

```
hold on;grid on
```

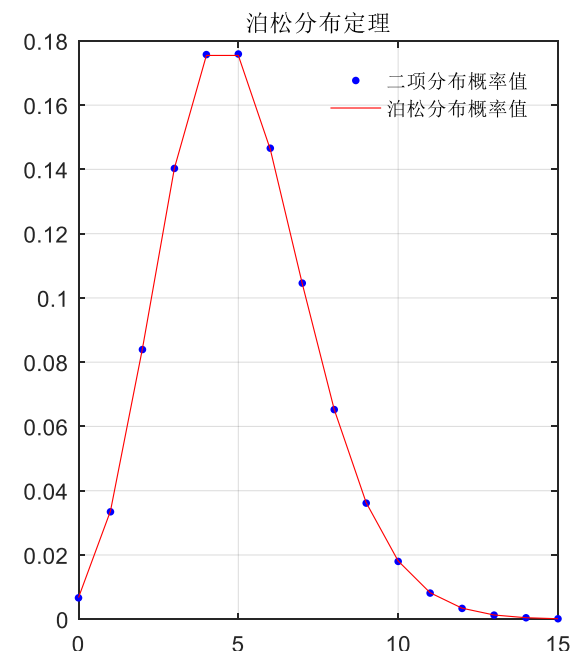
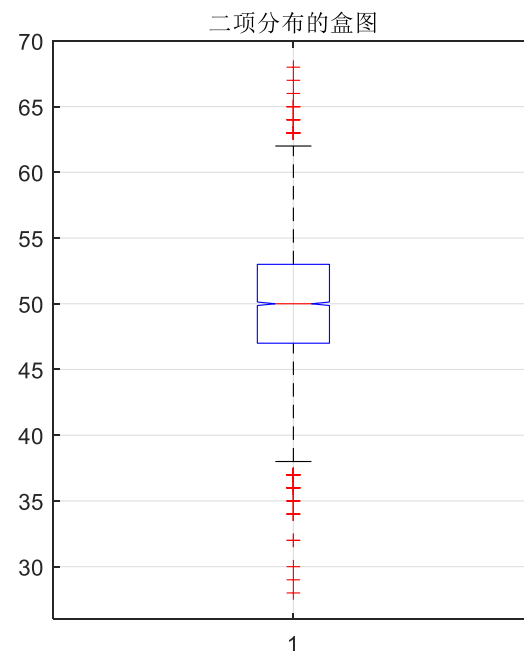
```
pp = poisspdf(k,5); plot(k,pp,'r-')
```

```
title('泊松分布定理')
```

```
legend('二项分布概率值','泊松分布概率值'); legend('boxoff')
```

泊松定理：若随机变量 $X \sim B(n, p)$ ，则当 $n$ 比较大， $p$ 比较小时，令 $\lambda = np$ ，则有

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}$$



## 4. 正态分布随机数及其检验

```
R = normrnd(0,1,100); %随机生成标准正态分布数据100*100
```

```
subplot(1,2,1); hist(R)
```

```
title('标准正态分布数据图hist')
```

```
subplot(1,2,2); histfit(R(:))
```

```
title('标准正态分布数据图histfit')
```

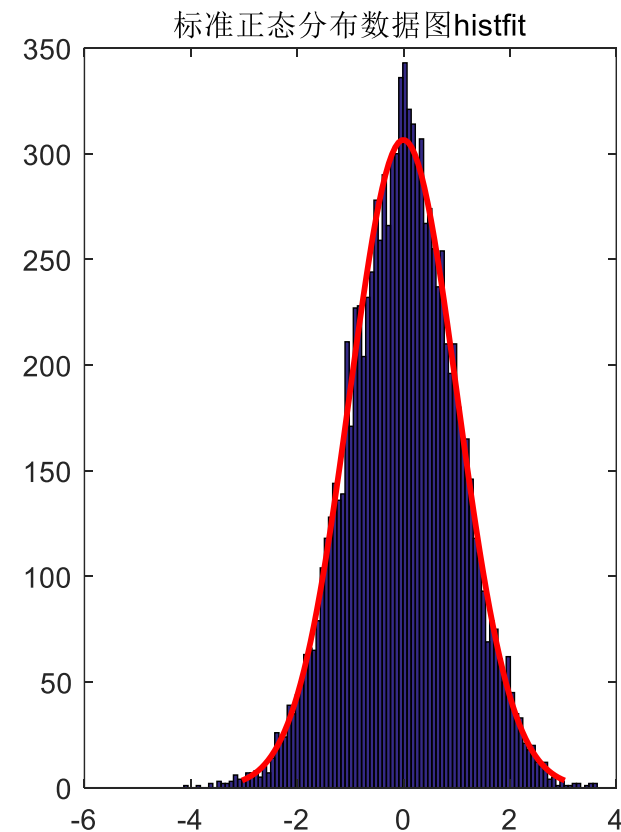
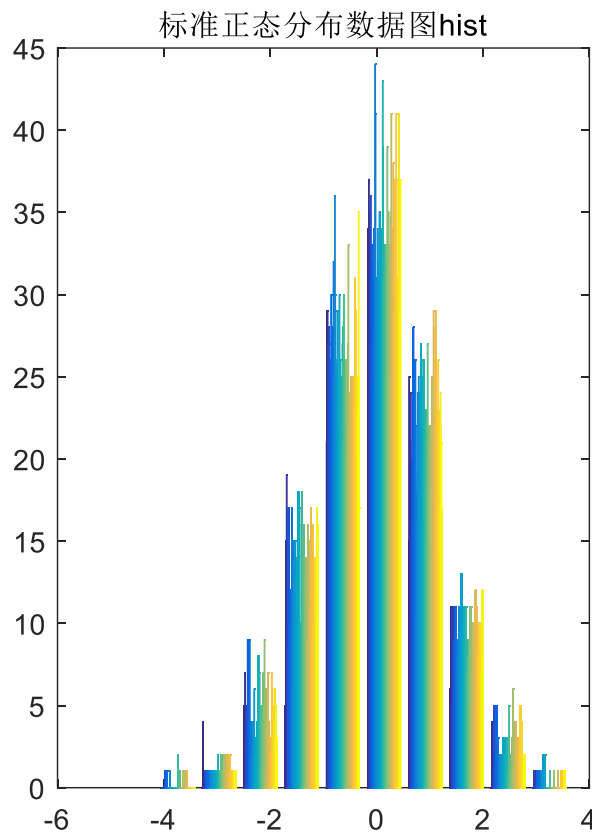
%随机抽取均值为50kg，方差为2的100个样本

```
R = normrnd(50,2,100,1);
```

%T检验均值和卡方检验方差

```
[h,p,muci,stats] = ttest(R,50,0.05)
```

```
[h,p,varci,stats] = vartest(R,4,0.05,0)
```



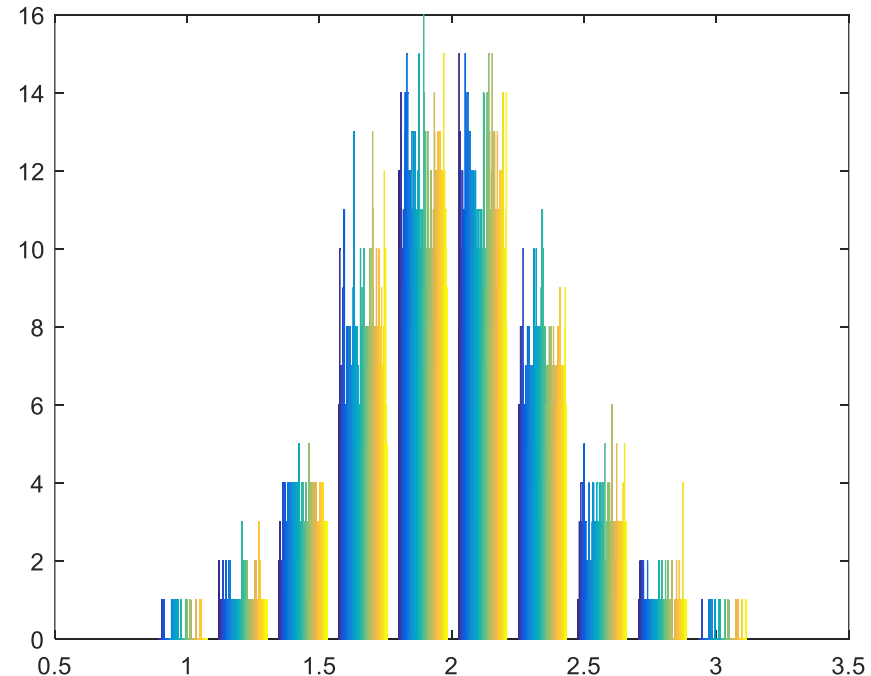
## 5. 通用函数求各分布的随机数据

### 求指定分布的随机数，函数`random`

- 格式 `y = random('name',A1,A2,A3,m,n)` %name的取值见常见分布函数表；A1, A2, A3为分布的参数；m, n指定随机数的行和列。

例：产生12000（30行400列）个均值为2，标准差为0.3的正态分布随机数。

```
>> Rnorm = random('norm',2,0.3,30,400);  
>> hist(Rnorm)
```





# 6. 常见分布函数表

name的取值	函数说明	name的取值	函数说明
'beta'或'Beta'	Beta分布	'ncf'或'Noncentral F'	非中心F分布
'bino'或'Binomial'	二项分布	'nct'或'Noncentral t'	非中心t分布
'chi2'或'Chisquare'	卡方分布	'ncx2'或'Noncentral Chi-square'	非中心卡方分布
'exp'或'Exponential'	指数分布	'norm'或'Normal'	正态分布
'f'或'F'	F分布	'poiss'或'Poisson'	泊松分布
'gam'或'Gamma'	GAMMA分布	'rayl'或'Rayleigh'	瑞利分布
'geo'或'Geometric'	几何分布	't'或'T'	T分布
'hyge'或'Hypergeometric'	超几何分布	'unif'或'Uniform'	均匀分布
'logn'或'Lognormal'	对数正态分布	'unid'或'Discrete Uniform'	离散均匀分布
'nbin'或'Negative Binomial'	负二项式分布	'weib'或'Weibull'	Weibull分布



---

# 感谢聆听

---