MATH 425 12/2/2022

Recall $\Gamma(n) = (n-1)\Gamma(n-1)$, $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$

<u>Example</u>: (a) Calculate $\Gamma(6)$.

(b) Calculate $\Gamma(7/2)$.

<u>Solution</u>: (a) If $k \in \mathbb{N}$ then $\Gamma(k) = (k-1)!$

$$\Gamma(6) = 5! = \underline{\underline{120}}$$

(b) $\Gamma(7/2) = 5/2 \cdot \Gamma(5/2) = 5/2 \cdot 3/2 \cdot \Gamma(3/2) = 5/2 \cdot 3/2 \cdot 1/2 \cdot \underbrace{\Gamma(1/2)}_{\sqrt{\pi}}$

$$= \boxed{\frac{15\sqrt{\pi}}{8}}$$

The χ^2 test. We said that if Z is a standard normal variable then Z^2 is distributed as $\Gamma_{1/2,1/2}$. Therefore if $Z_1, \dots, Z_k$ are independent standard normal variables then $Z_1^2 + Z_2^2 + \dots + Z_k^2$ is distributed as $\Gamma_{k/2,1/2}$. This is called χ^2_k or χ^2 with k degrees of freedom.

Pearson χ^2 test (large sample theory). We have a discrete normal variable distribution with n values, probability mass function $p_1, \dots, p_n$ ($p_1 + \dots + p_n = 1$). We have a random variable X and want to test whether it has this distribution.

(1924 - revolutionised the credibility of scientific research)

You perform N trials (independent)

O_i = # of times value i occurred $(i = 1, \dots, n)$

$E_i = p_i N$ = # of times value i is expected.

$$\boxed{\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \text{ has distribution } \chi^2_{n-1} = \Gamma_{(n-1)/2,\, 1/2}}$$

Example: A life insurance company has a table which says that

E_i	
10	10% die by the age of 60
10	20% die by the age of 70
10	30% die by the age of 75
20	50% die by the age of 80
50	

A young actuary joins the company and investigates death rates of their clients.
He takes a sample of 100 clients:

15 died before 60
15 died between 60 and 70
10 died between 70 and 75
10 died between 75 and 80
50 died over 80.

Can the actuary tell with 95% certainty that the company's table was wrong?

	O_i	E_i
<60	15	10
60<D<70	15	10
70<D<75	10	10
75<D<80	10	20
80<D	50	50

$$\sum_{i=1}^{5} \frac{(O_i - E_i)^2}{E_i} = \frac{5^2}{10} + \frac{5^2}{10} + \frac{0}{10} + \frac{10^2}{20} + \frac{0}{50} =$$

$$= 10$$

$$\chi_4^2 = 10 > \underbrace{9.488}_{95\text{ percentile}}$$

$$\boxed{\text{YES, THE TABLE OF THE COMPANY IS 95\% CERTAIN TO BE WRONG.}}$$

Another common statistical test is the t-test (Student test)

assumed name of a statistician working for Guinness

Small sample theory

X = normal variable with expectation μ

(but we do not know the standard deviation)

$X_1, \ldots, X_n$ are independent trials. Null hypothesis is that X_i all have the distribution X.

"general population"

Calculating the t test (with $n-1$ degrees of freedom) $(n>1)$

$$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$$

$$T_{n-1} = \frac{\bar{X} - \mu}{\sqrt{\dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n(n-1)}}}$$

T_k is distributed as

$$\frac{Z}{\sqrt{\frac{S}{k}}}$$

where Z, S are independent,

Z is standard normal

$$S = \chi^2_k$$

higher > average

95% ← confidence

not average 95%

density:

$$\frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1+\frac{t^2}{k}\right)^{-\frac{k+1}{2}}$$

Example: The quality of hops is measured on a scale of 0-30 with an average of 15. We got 4 samples which scored 10, 20, 20, 30. Is this supplier better than average with 95% confidence?

Solution: $\bar{X} = \frac{1}{4}(X_1 + X_2 + X_3 + X_4) = 20$

$$T = \frac{\bar{X} - 15}{\sqrt{\frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + (X_3-\bar{X})^2 + (X_4-\bar{X})^2}{4\cdot 3}}} = \frac{5}{\sqrt{\frac{10^2+0+0+10^2}{12}}} = \frac{5}{\sqrt{\frac{200}{12}}} \sim \underline{1.22}$$

critical 95%

T_3 one-tailed

2.35 $\boxed{\text{NO, we cannot tell}}$

(HW) (4) We cast a standard 6-sided die 60 times. Here is the table of how often numbers came up:

number	how many times
1	5
2	5
3	20
4	15
5	15
6	0

Do we know with 95% confidence that the die was biased?

(5) On a test, the average score is 80. Three students used a ~~test~~ prep company and scored 85, 90, 90. Do we know with 95% confidence that this company's preparation helps?