M358K: November 11th, 2022.

χ^2-distribution.

The following definition of the χ^2-distributed random variable can be extended.

For our purposes:

Let $Z_1, Z_2, \dots, Z_r$ be independent, standard normal r.v.s.

Define:

$$\boxed{X = Z_1^2 + Z_2^2 + \dots + Z_r^2}$$

We say that X has the χ^2-distribution w/ r degrees of freedom (parameter).

We write: $X \sim \chi^2(df = r)$

ν "nu"

Example. Let $X \sim \chi^2(df = 5)$.

Find $\mathbb{P}[1.145 \le X \le 12.83] = ?$

$\rightarrow$: $\mathbb{P}[X \le 12.83] - \mathbb{P}[X \le 1.145] =$

$= F_X(12.83) - F_X(1.145)$

1st Tables: $0.975 - 0.05 = 0.925$

2nd R: pchisq(12.83, df=5) - pchisq(1.145, df=5)

= 0.9250188

Goodness-of-fit

We are studying a multinomial experiment, possibly w/ categorical descriptions of possible outcomes.

The possible outcomes will be categories which are:

mutually exclusive and exhaustive.

Represent the categories as events $A_1, A_2, \ldots, A_k$.

In our probabilistic model, the parameters are

$$p_1, p_2, \cdots, p_k$$

w/ $p_i = \mathbb{P}[A_i]$ for $i = 1, 2, \ldots, k$.

Note: $\boxed{p_1 + p_2 + \cdots + p_k = 1}$

Repeat the same multinomial experiment n times independently.

Let X_i denote the number of times the outcome i occurred, for $i = 1, 2, \ldots, k$.

Note: $\boxed{X_1 + X_2 + \cdots + X_k = n}$ ☆

For our test statistic:

$$\boxed{Q^2 := \sum_{i=1}^{k} \frac{(X_i - n \cdot p_i)^2}{n \cdot p_i} \approx \chi^2(df = k-1)}$$

Works well for $np_i \geq 5$ for all $i = 1, 2, \ldots, k$