Chi-Square Dist.

Suppose that S^2 is the sample variance from a collection of iid. $N(\mu, \sigma^2)$ data.

then $\boxed{\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}}$ reads as: " a Chi-Square distribution with n−1 degrees of freedom"

Properties: $X \sim \chi^2_{n-1}$

then. ① $\mathbb{E}(X) = n-1$

② $Var(X) = 2(n-1)$

③ (Confidence Interval)

* $\chi^2_{n-1,\alpha}$: the α quantile of χ^2_{n-1}

then $$1-\alpha = P\left(\chi^2_{n-1,\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{n-1,1-\frac{\alpha}{2}}\right)$$

$\Rightarrow$ $100(1-\alpha)\%$ Confidence Interval for σ^2

$$= \left[\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$

$Z_1 \cdots Z_K$: iid $N(0,1)$

$$Q = \sum_{i=1}^{K} Z_i^2 \sim \chi^2(k) \rightarrow \text{model parameter}$$

(the only one)

$k > 0$

"degrees of freedom"

~~Unlike~~

Chi-square distribution is used primarily in hypothesis testing, and it's rarely used to model natural phenomena...

$\boxed{\text{Chi-Squared Test}}$ → for goodness of fit ~~of~~

an observed distribution to a theoretical one.

observed data

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

χ_c^2 → Chi-square statistic

E_i → expected values

Applications of χ^2-test ⇒ it's a way to ~~compare~~ decide is the variation of ⟨observed data⟩ just due to chance → noise or due to ~~of~~ one of the variables that you're actually testing.

↓

⟨expected values⟩

Example. flip a coin 100 times. → got 62 heads
38 tails

is it just a chance?

or somthing wrong with the coin?

Null hypothesis: "There is no significant different
H_0 between the observed and expected frequencies."

Degrees of Freedom = # of all possible outcomes − 1

= 2 − 1 = 1.

(Significance ~~significant~~ level)
Critical Value = 100% − Confidence → usually use 95%

= 5%