Lecture 22

niceguy

March 8, 2023

1 Comparison between Distributions

1.1 Central Limit Theorem

- Sample $X_1, \ldots, X_n$
- IID, finite σ^2
- $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- As $n \to \infty$, distribution of $\overline{X}$ tends to a normal distribution
- if X_i is normal, then $\overline{X}$ is normal $\forall n$

1.2 t distibution

- Sample $X_1, \ldots, X_n$
- IID, normal
- We do not need to know σ^2, but it is defined as

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

- T as defined below has normal distribution

$$T = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- If $n \geq 30, S \approx \sigma$, use Central Limit Theorem

1.3 χ^2 distribution

- Same as t distribution, where we assume normal distribution for X_i
- $$\chi^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

2 Quantiles

Given sample data $x_1, \ldots, x_n$, we have $q(f)$ where f is the fraction of data $\leq q(f)$. One can then plot $q(f)$ vs f.

Example 2.1 (Quantile Plot)**.** Given the data $-2, 0, 0, 1, 3, 3, 3, 4, 6$, $q(0.5) = 3$.

In general, $q(0.5)$ is called the sample median, $q(0.25$ the lower quartile and $q(0.75)$ the upper quartile.
This is the inverse of the Cumulative Distribution Function $F(x)$!

$$q(f) \approx \mu + \sigma\left(4.91\left(f^{0.14} - (1-f)^{0.14}\right)\right)$$

$q(f)$ can be approximated for a normal distribution.