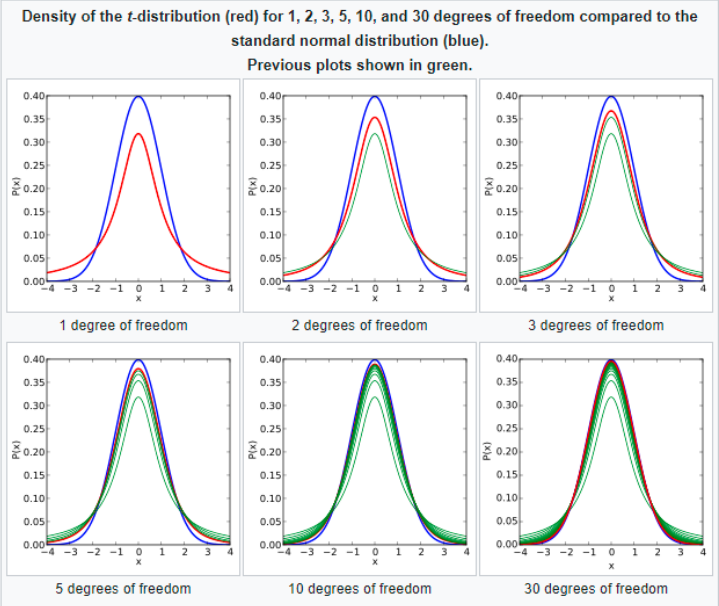
1. What is Student's t distribution? How is it defined? What is its pdf? What about the characteristics?

The T distribution, like the normal distribution, is bell-shaped and symmetric, but it has heavier tails, which means it tends to produce values that fall far from its mean. The T-distribution should only be used when population standard deviation is not known.

The larger the sample size, the more the t distribution looks like the normal distribution. In fact, for sample sizes larger than 20 (e.g. more degrees of freedom), the distribution is almost exactly like the normal distribution

The T Distribution (and the associated t scores), are used in hypothesis testing when you want to figure out if you should accept or reject the null hypothesis.

For calculating of t distribution we need the "degrees of freedom" that is just the sample size minus one.



Reference,

https://www.statisticshowto.com/probability-and-statistics/t-distribution/

https://www.investopedia.com/terms/t/tdistribution.asp

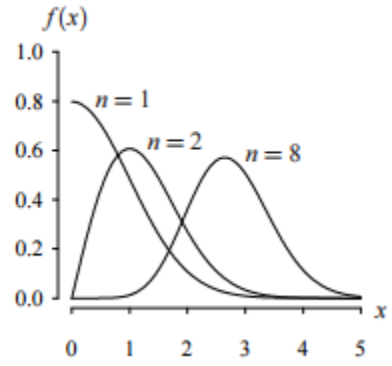
https://en.wikipedia.org/wiki/Student%27s_t-distribution

2. What is chi-square distribution? How is it defined? What is its pdf? What about the characteristics?

The chi-square distribution is a special case of the gamma distribution; A chi square distribution with n degrees of freedom is equal to a gamma distribution with a = n / 2 and b = 0.5 (or β = 2).

Let's say you have a random sample taken from a normal distribution. The chi square distribution is the distribution of the sum of these random samples squared . The degrees of freedom (k) are equal to the number of samples being summed. For example, if you have taken 10 samples from the normal distribution, then df = 10. The degrees of freedom in a chi square distribution is also its mean. Chi square distributions are always right skewed. However, the greater the degrees of freedom, the more the chi square distribution looks like a normal distribution.

$f(x) = 2^{(1-n/2)}\, x^{(n-1)}\, e^{(-(x_2)/2)} / \Gamma(n/2)$



There are two types of chi-square tests. Both use the chi-square statistic and distribution for different purposes:

- A chi-square goodness of fit test determines if sample data matches a populatio.
- A chi-square test for independence compares two variables in a contingency table to see if they are related.

Reference,

https://www.statisticshowto.com/probability-and-statistics/chi-square/

3. What is the difference between the word's "estimation" and "prediction"? Make research on Prediction Intervals for unknown population parameters. Are 95% confidence interval and 95% prediction interval for a parameter exactly the same? If not, what is the difference?

There is a sharp distinction between them in the standard model of a statistical problem. An **estimator** uses data to guess at a parameter while a **predictor** uses the data to guess at some random value that is not part of the dataset. We can tell that, an estimate is a guess about the true state of nature, but a prediction is a guess about another random value.

Note especially that this prediction has two separate sources of uncertainty: uncertainty in the data (xi,yi) leads to uncertainty in the estimated slope, intercept, and residual standard deviation (σ); in addition, there is uncertainty in just what value of Y(x) will occur. This additional uncertainty--because Y(x) is random--characterizes predictions. A prediction may look like an estimate (after all, α^+β^x estimates α+βx :-) and may even have the very same mathematical formula (p(x) can sometimes be the same as t(x)), but it will come with a greater amount of uncertainty than the estimate.

Distinguish,

- *purpose*: an estimator seeks to know a property of the true state of nature
- *uncertainty*: a predictor usually has larger uncertainty than a related estimator

the value of prediction interval and confidence interval generally are different in the below, you will find the definition of each one,

- Prediction interval,

In statistical inference, specifically predictive inference, a prediction interval is an estimate of an interval in which a future observation will fall, with a certain probability, given what has already been observed. Prediction intervals are often used in regression analysis.

- confidence interval

In frequentist statistics, a confidence interval (CI) is a range for estimating an unknown parameter. It can be bounded from one or both sides. A confidence interval is computed at a designated confidence level. The 95% level is most common, but other levels (such as 90% or 99%) are sometimes used.

In summary, with above information, we should consider the basic difference between estimation and prediction and usage of each. As you understand, the confidence interval is related to estimation and prediction interval is related to prediction.

Reference,

https://en.wikipedia.org/wiki/Confidence_interval

https://en.wikipedia.org/wiki/Prediction_interval

https://stats.stackexchange.com/questions/17773/what-is-the-difference-between-estimation-and-prediction#:~:text=Estimation%20is%20the%20calibration%20of,guessing%22%20of%20a%20future%20observation.