Econometric Analysis of Cross Section and Panel Data

Lecture 4: Normal Regression

Zhian Hu

Central University of Finance and Economics

Fall 2024

This Lecture

- Hansen (2022): Chapter 5
- This chapter introduces the normal regression model, which is a special case of the linear regression model.
- It is important as normality allows precise distributional characterizations and sharp inferences.
- Therefore in this chapter we introduce likelihood methods.

The Normal Distribution

- We say that a random variable Z has the standard normal distribution, or Gaussian, written $Z \sim \mathrm{N}(0,1)$, if it has the density

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad -\infty < x < \infty$$

- Properties:
  1. All integer moments of Z are finite.
  2. All odd moments of Z equal 0 .
  3. For any positive integer m

$$\mathbb{E}\left[Z^{2m}\right] = (2m-1)!! = (2m-1) \times (2m-3) \times \cdots \times 1$$

The Nornal Distribution

- If $Z \sim \mathrm{N}(0,1)$ and $X = \mu + \sigma Z$ for $\mu \in \mathbb{R}$ and $\sigma \geq 0$ then X has the univariate normal distribution, written $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$. By change-of-variables X has the density
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty$$
- The expectation and variance of X are μ and σ^2, respectively.

Multivariate Normal Distribution

- We say that the k-vector Z has a multivariate standard normal distribution, written $Z \sim \mathrm{N}(0, \boldsymbol{I}_k)$, if it has the joint density

$$f(x) = \frac{1}{(2\pi)^{k/2}} \exp\left(-\frac{x'x}{2}\right), \quad x \in \mathbb{R}^k$$

- The mean and covariance matrix of Z are 0 and $\boldsymbol{I}_k$, respectively.

Multivariate Normal Distribution

- If $Z \sim \mathrm{N}(0, \boldsymbol{I}_k)$ and $X = \mu + \boldsymbol{B}Z$ then the k-vector X has a multivariate normal distribution, written $X \sim \mathrm{N}(\mu, \Sigma)$ where $\Sigma = \boldsymbol{B}\boldsymbol{B}' \geq 0$. If $\Sigma > 0$ then by change-of-variables X has the joint density function

$$f(x) = \frac{1}{(2\pi)^{k/2}\det(\Sigma)^{1/2}} \exp\left(-\frac{(x-\mu)'\Sigma^{-1}(x-\mu)}{2}\right), \quad x \in \mathbb{R}^k$$

- The expectation and covariance matrix of X are μ and Σ, respectively.
- If $X \sim \mathrm{N}(\mu, \Sigma)$ and $Y = \boldsymbol{a} + \boldsymbol{B}X$, then $Y \sim \mathrm{N}(\boldsymbol{a} + \boldsymbol{B}\mu, \boldsymbol{B}\Sigma\boldsymbol{B}')$.

Properties of Multivariate Normal Distribution

- If (X, Y) are multivariate normal, X and Y are uncorrelated if and only if they are independent.
- If $X \sim \mathrm{N}(0, \boldsymbol{I}_k)$ then $X'X \sim \chi^2_k$, chi-square with k degrees of freedom.
- If $X \sim \mathrm{N}(0, \Sigma)$ with $\Sigma > 0$ then $X'\Sigma^{-1}X \sim \chi^2_k$ where $k = \dim(X)$.
- If $Z \sim \mathrm{N}(0, 1)$ and $Q \sim \chi^2_k$ are independent then $Z/\sqrt{Q/k} \sim t_k$, student t with k degrees of freedom.

卡方分布的作用：构造T统计量

Joint Normality and Linear Regression

- Suppose the variables (Y, X) are jointly normally distributed. Consider the best linear predictor of Y given X

$$Y = X'\beta + \alpha + e$$

最佳线性投影天然满足

  - So $\mathbb{E}[Xe] = 0$ and $\mathbb{E}[e] = 0$, so X and e are uncorrelated, and hence independent (Why?). because X and e are multivariate normal

- Independence implies that

$$\mathbb{E}[e \mid X] = \mathbb{E}[e] = 0 \quad \& \quad \mathbb{E}\left[e^2 \mid X\right] = \mathbb{E}\left[e^2\right] = \sigma^2$$

which are properties of a homoskedastic linear CEF.

- We have shown that when (Y, X) are jointly normally distributed, they satisfy a normal linear CEF

$$Y = X'\beta + \alpha + e, \quad e \sim \mathrm{N}\left(0, \sigma^2\right)$$

e is independent of X.

Normal Regression Model

- The normal regression model is the linear regression model with an independent normal error

$$Y = X'\beta + e$$
$$e \sim \mathrm{N}\left(0, \sigma^2\right)$$

- The normal regression model holds when (Y, X) are jointly normally distributed.
- For notational convenience, X contains the intercept.

Normal Regression Model

- The normal regression model implies that the conditional density of Y given X takes the form $f(y \mid x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{1}{2\sigma^2}(y - x'\beta)^2\right)$
- Under the assumption that the observations are mutually independent this implies that the conditional density of $(Y_1, \ldots, Y_n)$ given $(X_1, \ldots, X_n)$ is

$$
\begin{aligned}
f(y_1, \ldots, y_n \mid x_1, \ldots, x_n) &= \prod_{i=1}^{n} f(y_i \mid x_i) \\
&= \prod_{i=1}^{n} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - x_i'\beta)^2\right) \\
&= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2\right) \\
&\stackrel{\text{def}}{=} L_n(\beta, \sigma^2)
\end{aligned}
$$

- This is called the likelihood function when evaluated at the sample data.

Normal Regression Model

- For convenience it is typical to work with the natural logarithm

$$\log L_n\left(\beta, \sigma^2\right) = -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(Y_i - X_i'\beta\right)^2 \overset{\text{def}}{=} \ell_n\left(\beta, \sigma^2\right)$$

  which is called the log-likelihood function.
- The maximum likelihood estimator (MLE) $(\widehat{\beta}_{\text{mle}}, \widehat{\sigma}^2_{\text{mle}})$ is the value which maximizes the log-likelihood.
- We can write the maximization problem as

$$\left(\widehat{\beta}_{\text{mle}}, \widehat{\sigma}^2_{\text{mle}}\right) = \underset{\beta\in\mathbb{R}^k, \sigma^2>0}{\operatorname{argmax}} \ \ell_n\left(\beta, \sigma^2\right)$$

Normal Regression Model

黑塞矩阵

- The maximizers $\left(\widehat{\beta}_{\text{mle}}, \widehat{\sigma}^2_{\text{mle}}\right)$ jointly solve the first-order conditions (FOC)

$$0 = \frac{\partial}{\partial \beta} \ell_n\left(\beta, \sigma^2\right)\Big|_{\beta=\widehat{\beta}_{\text{mle}}, \sigma^2=\widehat{\sigma}^2_{\text{mle}}} = \frac{1}{\widehat{\sigma}^2_{\text{mle}}} \sum_{i=1}^{n} X_i \left(Y_i - X_i' \widehat{\beta}_{\mathrm{mle}}\right)$$

$$0 = \frac{\partial}{\partial \sigma^2} \ell_n\left(\beta, \sigma^2\right)\Big|_{\beta=\widehat{\beta}_{\mathrm{mle}}, \sigma^2=\widehat{\sigma}^2_{\text{mle}}} = -\frac{n}{2\widehat{\sigma}^2_{\text{mle}}} + \frac{1}{2\widehat{\sigma}^4_{\text{mle}}} \sum_{i=1}^{n} \left(Y_i - X_i' \widehat{\beta}_{\text{mle}}\right)^2$$

- The first FOC is proportional to the first-order conditions for the least squares minimization problem. It follows that the MLE satisfies

$$\widehat{\beta}_{\text{mle}} = \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} X_i Y_i\right) = \widehat{\beta}_{\text{ols}}$$

- Solving the second FOC for $\widehat{\sigma}^2_{\text{mle}}$ we find

$$\widehat{\sigma}^2_{\mathrm{mle}} = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - X_i' \widehat{\beta}_{\mathrm{mle}}\right)^2 = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - X_i' \widehat{\beta}_{\mathrm{ols}}\right)^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{e}_i^2 = \widehat{\sigma}^2_{\mathrm{ols}}$$

Distribution of OLS Coefficient Vector

- In the normal linear regression model we can derive exact sampling distributions for the OLS/MLE estimator, residuals, and variance estimator.
- The normality assumption $e \mid X \sim \mathrm{N}\left(0, \sigma^2\right)$ combined with independence of the observations has the multivariate implication

$$\boldsymbol{e} \mid \boldsymbol{X} \sim \mathrm{N}\left(0, \boldsymbol{I}_n \sigma^2\right)$$

- That is, the error vector $\boldsymbol{e}$ is independent of $\boldsymbol{X}$ and is normally distributed.
- Recall that the OLS estimator satisfies

$$\widehat{\beta} - \beta = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \boldsymbol{X}'\boldsymbol{e}$$

β^=(X'X)^{-1}X'(Xβ+e)

which is a linear function of $\boldsymbol{e}$.

Distribution of OLS Coefficient Vector

- Since linear functions of normals are also normal this implies that conditional on $\boldsymbol{X}$

$$\begin{aligned}\widehat{\beta}-\beta \mid \boldsymbol{X} &\sim \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\mathrm{N}\left(0, \boldsymbol{I}_n\sigma^2\right) \\ &\sim \mathrm{N}\left(0, \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\right) \\ &= \mathrm{N}\left(0, \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\right)\end{aligned}$$

- In the normal regression model,

$$\widehat{\beta} \mid \boldsymbol{X} \sim \mathrm{N}\left(\beta, \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\right)$$

- Letting β_j and $\widehat{\beta}_j$ denote the j^{th} elements of β and $\widehat{\beta}$, we have

$$\widehat{\beta}_j \mid \boldsymbol{X} \sim \mathrm{N}\left(\beta_j, \sigma^2\left[\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\right]_{jj}\right)$$

Distribution of OLS Residual Vector

- Recall that $\widehat{\boldsymbol{e}} = \boldsymbol{Me}$ where $\boldsymbol{M} = \boldsymbol{I}_n - \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'$. So conditional on $\boldsymbol{X}$

$$\widehat{\boldsymbol{e}} = \boldsymbol{Me} \mid \boldsymbol{X} \sim \mathrm{N}\left(0, \sigma^2 \boldsymbol{MM}\right) = \mathrm{N}\left(0, \sigma^2 \boldsymbol{M}\right)$$

- Furthermore, it is useful to find the joint distribution of β and $\widehat{\boldsymbol{e}}$.

$$\begin{pmatrix} \widehat{\beta} - \beta \\ \widehat{\boldsymbol{e}} \end{pmatrix} = \begin{pmatrix} \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{e} \\ \boldsymbol{Me} \end{pmatrix} = \begin{pmatrix} \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}' \\ \boldsymbol{M} \end{pmatrix} \boldsymbol{e}$$

- The vector has a joint normal distribution with covariance matrix

BB'\sigma^2

$$\begin{pmatrix} \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} & 0 \\ 0 & \sigma^2 \boldsymbol{M} \end{pmatrix}$$

XM=0,因为正交矩阵

协方差为0

- Since the off-diagonal block is zero it follows that $\widehat{\beta}$ and $\widehat{\boldsymbol{e}}$ are statistically independent.

Distribution of Variance Estimator

- Next, consider the variance estimator s^2.
- It satisfies $(n-k)\,s^2 = \widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}} = \boldsymbol{e}'\boldsymbol{M}\boldsymbol{e}$. The spectral decomposition of $\boldsymbol{M}$ is $\boldsymbol{M} = \boldsymbol{H}\Lambda\boldsymbol{H}'$ where $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_n$ and Λ is diagonal with the eigenvalues of $\boldsymbol{M}$ on the diagonal.
- Since $\boldsymbol{M}$ is idempotent with rank $n-k$, it has $n-k$ eigenvalues equalling 1 and k eigenvalues equalling 0 , so

$$\Lambda = \begin{bmatrix} \boldsymbol{I}_{n-k} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_k \end{bmatrix}$$

Distribution of Variance Estimator

- Let $\boldsymbol{u} = \boldsymbol{H}'\boldsymbol{e} \sim \mathrm{N}\left(\mathbf{0}, \boldsymbol{I}_n\sigma^2\right)$ and partition $\boldsymbol{u} = (\boldsymbol{u}_1', \boldsymbol{u}_2')'$ where $\boldsymbol{u}_1 \sim \mathrm{N}\left(0, \boldsymbol{I}_{n-k}\sigma^2\right)$. Then

$$
\begin{aligned}
(n-k)s^2 &= \boldsymbol{e}'\boldsymbol{M}\boldsymbol{e} \\
&= \boldsymbol{e}'\boldsymbol{H}\begin{bmatrix} \boldsymbol{I}_{n-k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\boldsymbol{H}'\boldsymbol{e} \\
&= \boldsymbol{u}'\begin{bmatrix} \boldsymbol{I}_{n-k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\boldsymbol{u} \\
&= \boldsymbol{u}_1'\boldsymbol{u}_1 \\
&\sim \sigma^2\chi^2_{n-k}
\end{aligned}
$$

- We see that in the normal regression model the exact distribution of s^2 is a scaled chi-square. Since $\widehat{\boldsymbol{e}}$ is independent of $\widehat{\beta}$ it follows that s^2 is independent of $\widehat{\beta}$ as well.

t-statistic

- We already know that $\widehat{\beta}_j \mid \boldsymbol{X} \sim \mathrm{N}\left(\beta_j, \sigma^2 \left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\right]_{jj}\right)$. So

$$\frac{\widehat{\beta}_j - \beta_j}{\sqrt{\sigma^2 \left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\right]_{jj}}} \sim \mathrm{N}(0,1)$$

- Now take the standardized statistic and replace the unknown variance σ^2 with its estimator s^2. We call this a t-ratio or t-statistic

$$T = \frac{\widehat{\beta}_j - \beta_j}{\sqrt{s^2 \left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\right]_{jj}}} = \frac{\widehat{\beta}_j - \beta_j}{s\left(\widehat{\beta}_j\right)}$$

where $s\left(\widehat{\beta}_j\right)$ is the classical (homoskedastic) standard error for $\widehat{\beta}_j$.

t-statistic

- With algebraic re-scaling we can write the t-statistic as the ratio of the standardized statistic and the square root of the scaled variance estimator.

$$
\begin{aligned}
T &= \frac{\widehat{\beta}_j - \beta_j}{\sqrt{\sigma^2 \left[\left(\boldsymbol{X}'\boldsymbol{X} \right)^{-1} \right]_{jj}}} \Bigg/ \sqrt{\frac{(n-k)s^2}{\sigma^2} / (n-k)} \\
&\sim \frac{\mathrm{N}(0,1)}{\sqrt{\chi^2_{n-k}/(n-k)}} \\
&\sim t_{n-k}
\end{aligned}
$$

  a student t distribution with $n-k$ degrees of freedom.

- This derivation shows that the t-ratio has a sampling distribution which depends only on the quantity $n-k$.

t-statistic

- An important caveat about the above theorem is that it only applies to the t-statistic constructed with the homoskedastic (old-fashioned) standard error.
- It does not apply to a t-statistic constructed with any of the heteroskedasticity-robust standard errors.
- In fact, the robust t-statistics can have finite sample distributions which deviate considerably from t_{n-k} even when the regression errors are independent $\mathrm{N}\left(0, \sigma^2\right)$.
- Thus the distributional result in the above theorem and the use of the t distribution in finite samples is only exact when applied to classical t-statistics under the normality assumption.

Confidence Intervals for Regression Coefficients

- The OLS estimator $\widehat{\beta}$ is a point estimator for a coefficient β.
- A broader concept is a set or interval estimator which takes the form $\widehat{C} = [\widehat{L}, \widehat{U}]$.
- The goal of an interval estimator $\widehat{C}$ is to contain the **true value**, e.g. $\beta \in \widehat{C}$, with high probability.
- The interval estimator $\widehat{C}$ is a function of the data and hence is random.

Confidence Intervals for Regression Coefficients

- An interval estimator $\widehat{C}$ is called a $1-\alpha$ confidence interval when $\mathbb{P}[\beta \in \widehat{C}] = 1-\alpha$ for a selected value of α.
- The value $1-\alpha$ is called the coverage probability. Typical choices for the coverage probability $1-\alpha$ are 0.95 or 0.90.
- The probability calculation $\mathbb{P}[\beta \in \widehat{C}]$ is easily mis-interpreted as treating β as random and $\widehat{C}$ as fixed. (The probability that β is in $\widehat{C}$.)
- This is not the appropriate interpretation. Instead, the correct interpretation is that the probability $\mathbb{P}[\beta \in \widehat{C}]$ treats the point β as fixed and the set $\widehat{C}$ as random. It is the probability that the random set $\widehat{C}$ covers (or contains) **the fixed true coefficient** β.

Confidence Intervals for Regression Coefficients

- A good choice for a confidence interval for the regression coefficient β is obtained by adding and subtracting from the estimator $\widehat{\beta}$ a fixed multiple of its standard error:

$$\widehat{C} = [\widehat{\beta} - c \times s(\widehat{\beta}), \quad \widehat{\beta} + c \times s(\widehat{\beta})]$$

  where $c > 0$ is a pre-specified constant which determines the coverage probability.
- This confidence interval is symmetric about the point estimator $\widehat{\beta}$ and its length is proportional to the standard error $s(\widehat{\beta})$.

Confidence Intervals for Regression Coefficients

- Equivalently, $\widehat{C}$ is the set of parameter values for β such that the t-statistic $T(\beta)$ is smaller (in absolute value) than c, that is

$$\widehat{C} = \{\beta : |T(\beta)| \leq c\} = \left\{\beta : -c \leq \frac{\widehat{\beta} - \beta}{s(\widehat{\beta})} \leq c\right\}$$

- The coverage probability of this confidence interval is

$$\begin{aligned} \mathbb{P}[\beta \in \widehat{C}] &= \mathbb{P}[|T(\beta)| \leq c] \\ &= \mathbb{P}[-c \leq T(\beta) \leq c] \end{aligned}$$

- Since the t-statistic $T(\beta)$ has the t_{n-k} distribution, it equals $F(c) - F(-c)$, where $F(u)$ is the student t distribution function with $n - k$ degrees of freedom.
- Since $F(-c) = 1 - F(c)$, we can write it as

$$\mathbb{P}[\beta \in \widehat{C}] = 2F(c) - 1$$

  This is the coverage probability of the interval $\widehat{C}$, and only depends on the constant c.

Confidence Intervals for Regression Coefficients

- When the degree of freedom is large the distinction between the student t and the normal distribution is negligible.
- In particular, for $n - k \geq 61$ we have $c \approx 2.00$ for a 95% interval.
- Using this value we obtain the most commonly used confidence interval in applied econometric practice:
$$\widehat{C} = [\widehat{\beta} - 2s(\widehat{\beta}), \quad \widehat{\beta} + 2s(\widehat{\beta})]$$
- This is a useful rule-of-thumb. This 95% confidence interval $\widehat{C}$ is simple to compute and can be easily calculated from coefficient estimates and standard errors.

t Test

- A typical goal in an econometric exercise is to assess whether or not a coefficient β equals a specific value β_0.
- Often the specific value to be tested is $\beta_0 = 0$ but this is not essential. This is called **hypothesis testing**.
- For simplicity write the coefficient to be tested as β. The null hypothesis is

$$\mathbb{H}_0 : \beta = \beta_0$$

- This states that the hypothesis is that the true value of β equals the hypothesized value β_0.
- The alternative hypothesis is the complement of $\mathbb{H}_0$, and is written as

$$\mathbb{H}_1 : \beta \neq \beta_0$$

t Test

- We are interested in testing $\mathbb{H}_0$ against $\mathbb{H}_1$.
- The method is to design a statistic which is informative about $\mathbb{H}_1$ and to characterize its sampling distribution.
- The standard statistic is the absolute value of the t-statistic

$$|T| = \left| \frac{\widehat{\beta} - \beta_0}{s(\widehat{\beta})} \right|$$

- If $\mathbb{H}_0$ is true then we expect $|T|$ to be small, but if $\mathbb{H}_1$ is true then we would expect $|T|$ to be large.
- Hence the standard rule is to reject $\mathbb{H}_0$ in favor of $\mathbb{H}_1$ for large values of the t-statistic $|T|$ and otherwise fail to reject $\mathbb{H}_0$. Thus the hypothesis test takes the form: Reject $\mathbb{H}_0$ if $|T| > c$.

t Test

- The constant c which appears in the statement of the test is called the critical value.

$$\begin{aligned}\mathbb{P}\left[\text{ Reject } \mathbb{H}_0 \mid \mathbb{H}_0\right] &= \mathbb{P}\left[|T| > c \mid \mathbb{H}_0\right] \\ &= \mathbb{P}\left[T > c \mid \mathbb{H}_0\right] + \mathbb{P}\left[T < -c \mid \mathbb{H}_0\right] \\ &= 1 - F(c) + F(-c) \\ &= 2(1 - F(c))\end{aligned}$$

- We select the value c so that this probability equals a pre-selected value called the significance level which is typically written as α.
- It is conventional to set $\alpha = 0.05$, though this is not a hard rule. We then select c so that $F(c) = 1 - \alpha/2$, which means that c is the $1 - \alpha/2$ quantile (inverse CDF) of the t_{n-k} distribution.
- With this choice the decision rule "Reject $\mathbb{H}_0$ if $|T| > c$ " has a significance level (false rejection probability) of α.

t Test

- A simplification of the above test is to report what is known as the **p**-value of the test.
- In general, when a test takes the form "Reject $\mathbb{H}_0$ if $S > c$" and S has null distribution $G(u)$ then the p-value of the test is $p = 1 - G(S)$.
- A test with significance level α can be restated as "Reject $\mathbb{H}_0$ if $p < \alpha$ ".
- It is sufficient to report the p-value p and we can interpret the value of p as indexing the test's strength of rejection of the null hypothesis.
- Thus a p-value of 0.07 might be interpreted as "nearly significant", 0.05 as "borderline significant", and 0.001 as "highly significant".
- In the context of the normal regression model the p-value of a t-statistic $|T|$ is $p = 2\,(1 - F_{n-k}(|T|)\ )$ where F_{n-k} is the t_{n-k} CDF.