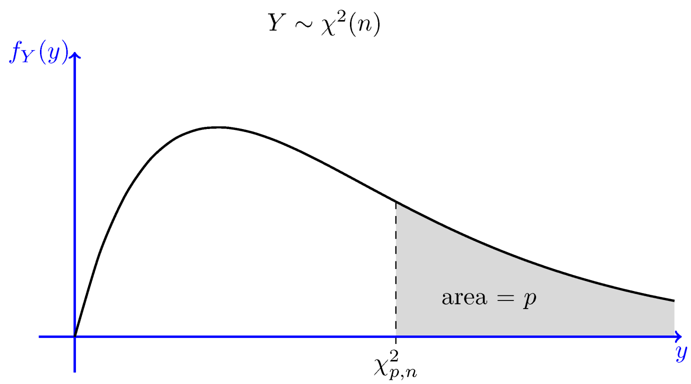

Figure 8.6 - The definition of $\chi^2_{p,n}$.

Now, why do we need the chi-squared distribution? One reason is the following theorem, which we will use in estimating the variance of normal random variables.

Theorem 8.3. Let X_1, X_2, $\cdots$, X_n be i.i.d. $N(\mu, \sigma)$ random variables. Also, let S^2 be the standard variance for this random sample. Then, the random variable Y defined as

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

has a chi-squared distribution with $n-1$ degrees of freedom, i.e., $Y \sim \chi^2(n-1)$. Moreover, $\overline{X}$ and S^2 are independent random variables.

The t-Distribution

The next distribution that we need is the **Student's t-distribution** (or simply the **t-distribution**). Here, we provide the definition and some properties of the t-distribution.

The t-Distribution

Definition 8.2. Let $Z \sim N(0,1)$, and $Y \sim \chi^2(n)$, where $n \in \mathbb{N}$. Also assume that Z and Y are independent. The random variable T defined as

$$T = \frac{Z}{\sqrt{Y/n}}$$

is said to have a t-distribution with n *degrees of freedom* shown by

$$T \sim T(n).$$

Properties:

1. The t-distribution has a bell-shaped PDF centered at 0, but its PDF is more spread out than the normal PDF (Figure 8.7).
2. $ET = 0$, for $n > 0$. But ET, is undefined for $n = 1$.
3. $\mathrm{Var}(T) = \frac{n}{n-2}$, for $n > 2$. But, $\mathrm{Var}(T)$ is undefined for $n = 1, 2$.
4. As n becomes large, the t density approaches the standard normal PDF. More formally, we can write

$$T(n) \xrightarrow{d} N(0,1).$$

5. For any $p \in [0,1]$ and $n \in \mathbb{N}$, we define $t_{p,n}$ as the real value for which

$$P(T > t_{p,n}) = p.$$

   Since the t-distribution has a symmetric PDF, we have

$$t_{1-p,n} = -t_{p,n}.$$

   In MATLAB, to compute $t_{p,n}$ you can use the following command: `tinv(1 - p, n)`.

Figure 8.7 shows the PDF of t-distribution for some values of n and compares them with the PDF of the standard normal distribution. As we see, the t density is more spread out than the standard normal PDF. Figure 8.8 shows $t_{p,n}$.

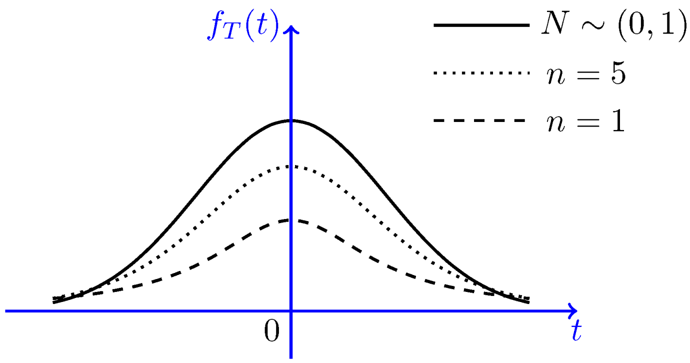

Figure 8.7 - The PDF of t-distribution for some values of n compared with the standard normal PDF.

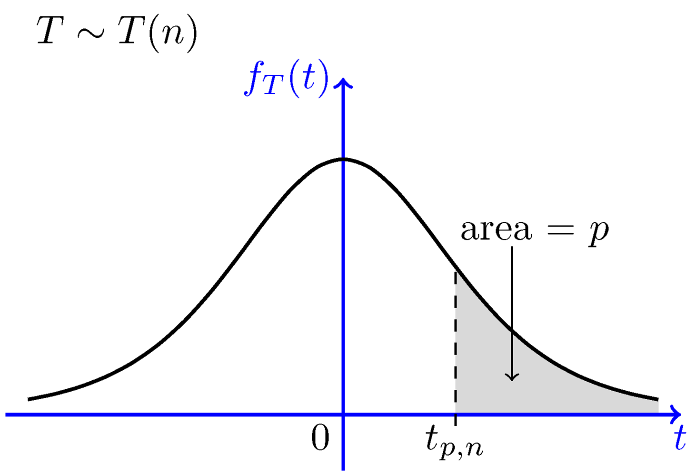

Figure 8.8 - The definition of $t_{p,n}$.

Why do we need the t-distribution? One reason is the following theorem which we will use in estimating the mean of normal random variables.