Formal statistical hypotheses testing is one of the topics in this class. As you know or will find out later, there are many details that tend to make a quick, complete understanding rather difficult. The following less formal discussion of the idea behind hypothesis testing and estimation provides some insight into what statistics is all about. The simple idea is this.

Probability tells us, under the assumption that our hypothesis is correct, the distribution of certain observable statistical quantities; if our observed values (from running an actual experiment) are too unbelievable in light of this knowledge from probability then we conclude our hypothesis was incorrect.

Example 1.

According to Larson and Marx (1985) the "cause and mode of transmission of leukemia remain largely unknown." They go on to say that "prevailing medical opinion holds that most forms of leukemia are not contagious." There is evidence that some forms, such as the childhood variety, may be. Events in Niles Illinois in the late fifties provide a possible example.

In the 5 and 1/3 year period from 1956 to the first four months of 1961 physicians in Niles Illinois, a Chicago suburb, reported a total of 8 cases of leukemia among children less than 15 years of age. The population of persons in this category numbered 7,076 in Niles while for all of Cook County, excluding Niles, there were during the same time period 1,152,695 children less than 15 years of age. Among those children there were 286 diagnosed cases of leukemia.

If persons in the at-risk category contract leukemia independently then, letting p be the probability that an individual child contracted the disease in the given time period, one would conclude that the number Y of cases in Niles is

$$Y = \sum_{j=1}^{7076} X_j \ ,$$

where

$$X_j = \begin{cases} 1 & \text{if jth child contracts leukemia} \\ 0 & \text{if the jth child does not} \end{cases} ,$$

and that Y should be an observation of a B(7076, p) random variable. Furthermore, if conditions in Niles were the same as those in the rest of Cook County and independence holds then an excellent estimate of p (for all practical purposes, p itself) is provided by the events in Cook County outside of Niles. We take therefore

$$p = \frac{286}{1152695} = 2.48 \times 10^{-4} \ .$$

Under independence and if Niles and the rest of Cook County were the same regarding the chances of contracting the disease, the probability of observing as many as 8 cases in Niles is precisely

$$P[\,Y \geq 8\,] = \sum_{j \geq 8} \binom{7076}{j} p^j (1-p)^{7076-j}\ .$$

Rather than attempting to calculate this we note that np = (7076) 2.48×10^{-4} = 1.75 and this is in the range of our Poisson tables. From our study of the Poisson approximation to the binomial we know that if W ~ P(1.75) then P[Y ≥ 8] ~ P[W ≥ 8] and consulting the Poisson tables we see that this is roughly 4.9×10^{-4} . This probability is so small that the event that happened is unbelievable. We conclude therefore that our assumptions are incorrect; either

(i) independence is violated (for example, leukemia is communicable or there are genetic factors)
or
(ii) conditions relating to probability of contracting the disease in Niles were different from those in the rest of Cook County,
or
(iii) some other factors cause the binomial assumption to be inappropriate. ◊

Example 2. (relationship between estimation and hypothesis testing)

Consulting the list of formal hypothesis tests one finds that if $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ then to test

$$H_0 : \mu = \mu_0 \text{ against } H_A : \mu \neq \mu_0$$

at level α reject H_0 if $|\,\tau\,| > t_{n-1,\alpha/2}$, where

$$\tau = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}\ .$$

This procedure can be shown mathematically to be optimal, but an heuristic understanding is sufficient for our study and is simply based upon the reasoning expressed above, the fact that if the null hypothesis were correct then the test statistic τ would have a t-distribution with n-1 degrees of freedom, and that evidence against the mean being the hypothesized value μ_0 is provided by values of $\overline{x}$ too far away from μ_0, either above or below.

You also know, or will before the successful completion of the course, that a $(1-\alpha)100\%$ confidence interval for μ is $(\overline{x} - \frac{s}{\sqrt{n}} t_{n-1,\alpha/2}, \overline{x} + \frac{s}{\sqrt{n}} t_{n-1,\alpha/2})$. It is a simple exercise in elementary algebra to show that for given observed values of $\overline{x}$ and s the confidence interval contains precisely all of the possible hypothesized μ_0's that would not have led to rejection of the null hypothesis. Perhaps this could be rephrased as: our interval estimate of the mean is the collection of mean values which are consistent with the data. ◊

Reference

Larsen, R. and Marx, M. *An Introduction to Probability and Its Applications.* (1985) Prentice Hall.