mixed experiment, $x$ has the mixed distribution

$$g(x; p) = pf(x; \theta = 0) + (1 - p)f(x; \theta = 1).$$

This mixed distribution can be used to estimate the proportion of ineffective compounds.

It is important to know that not all five statistical principles are mutually independent. The next paradox, the Birnbaum experiment, will show you why.

### 3.5.2 Paradox of the Birnbaum Experiment*

The argument from the Birnbaum Experiment purports to derive the strong likelihood principle (SLP) from sufficiency and conditionality principles (Birnbaum, 1965; Berger and Wolpert, 1988; Casella and Berger, 2002). Let’s outline Birnbaum’s arguments.

If two observations $y^*$ and $y''^*$ are from two experiments $E'$ and $E''$, respectively, and have proportional likelihoods, then they have a common sufficient statistic $T(y^*) = T(y''^*)$, since the likelihood itself is a sufficient statistic. Furthermore, the SLP states that the inference about parameter $\theta$ based on $y^*$ from $E'$ should be the same as that based on $y''^*$ from $E''$.

$$\Upsilon_{E'}(y^*) = \Upsilon_{E''}(y''^*),$$

where $\Upsilon$ means “inference.”

We now consider a mixture experiment wherein a coin (fair or biased) is flipped, with “heads” leading to performing $E'$ and reporting the outcome $y'$, and “tails” leading to $E''$ and reporting the outcome $y''$. Each outcome would have two components $(E^j, y^j)$ ($j = '$ or '$), and the distribution for the mixture would be sampled over the distinct sample spaces of $E'$ and $E''$.

In the Birnbaum experiment there are two possible cases in terms of the outcomes:

- **Case 1**, when $E'$ and $E''$ have “star paired” outcomes (i.e., their likelihoods are proportional), we define the test statistic using the common sufficient statistic $T(y^*)$. In other words, we always report $T(y^*)$, even if the data $y''^*$ are actually observed.
- **Case 2**, when $E'$ and $E''$ don’t have “star paired” outcomes (i.e., their likelihoods are not proportional), we formulate the test statistic (usually) based on the experiment performed $E^j$ and outcome $y^j$, where $j = '$ or '$.
That is, Birnbaum’s experiment, $E_{BB}$, is based on the statistic $T_{BB}$, given by

$$T_{BB}(E^j, y^j) = \begin{cases} T(y^*) & \text{if } j = 1 \text{ and } y' = y'' \text{ or if } j = 2 \text{ and } y'' = y'^* \\ T(E^j, y^j), & \text{otherwise} \end{cases}$$

Because the argument for the SLP is dependent on Case 1 outcomes, we may focus only on them for now. Let’s start with the following two premises:

1. Because there is a common sufficient statistic $T_{BB}(E^j, y^j) = T(y^*)$ for $E'$ and $E''$, based on the sufficiency principle, we have the first premise of the argument:

$$\Upsilon_{E_{BB}}(E', y'^*) = \Upsilon_{E_{BB}}(E'', y''^*). \tag{3.5}$$

2. The argument next points out that the conditionality principle tells us that, once it is known which of $E'$ or $E''$ produced the outcome, we should compute the inference just as if it were known all along that $E^j$ would be performed. Applying the conditionality principle to Birnbaum’s mixture gives the premise:

$$\Upsilon_{E_{BB}}(E^j, y^j^*) = \Upsilon_{E^j}(y^j^*). \tag{3.6}$$

Expanding (3.6) into two equations for $j = ' \text{ and } ''$ and using (3.5), we immediately have $\Upsilon_{E'}(y'^*) = \Upsilon_{E''}(y''^*)$, that is, the inference from $y'^*$ is identical to the inference from $y''^*$, which is the SLP.

Paradoxically, Professor of Philosophy Deborah Mayo (Mayo and Spanos, 2010, p. 305–314) claims to have disproved Birnbaum’s argument by pointing to an apparently hidden fault in the proof. Mayo states: “The problem is that, even by allowing all the premises to be true, the conclusion could “follow” only if it is assumed that you both should and should not use the conditional formulation. The antecedent of premise (1) is the denial of the antecedent of premise (2).”

However, Mayo’s disproof is faulty because her presumption about the antecedent of premise (1) in Birnbaum’s argument is odd. A sufficient statistic is sufficient for a FAMILY of distributions with different values of the parameter $\theta$; such a family of distributions often consists of the distributions under $H_o$ and $H_a$. Remember, given a sufficient statistic $T$, the distribution of $x$ will vary as $\theta$ varies. Therefore, her statement about the sufficient statistic under a mixed distribution (a fixed distribution) is irrelevant.
Frequentists don’t believe the likelihood principle because they don’t believe conditionality, as we discussed in Section 3.2.1.