23 Two stage procedures.

Suppose we are interested in constructing a 95% confidence interval for the mean μ of a normal populaton with an unknown variance σ^2, We want to make sure that the width of the interval is no more than 0.1. In other words we want to construct a random interval $(a-t, a+t)$ such that $|t| \leq .05$, and the probability $P\{\mu \in (a-t, a+t)\} \geq 0.95$ for all μ and σ^2. We are free to choose as many observations as we need to.

The usual procedure is to take a sample of size n and compute $\bar{x}$ and s, the sample meean and the sample standard deviation based on the n observations. Since

$$t_{n-1} = \frac{\bar{x}-\mu}{s}\sqrt{n-1}$$

has a t with $n-1$ degrees of freedom, we cwn determine from the tables a value t_0 such that

$$P\{|t_{n-1}| \leq t_0\} = 0.95$$

Then the interval $(\bar{x} - \frac{st_0}{n-1}, \bar{x} + \frac{st_0}{n-1})$ provides a confidence interval of width $\frac{2st_0}{n-1}$ for the mean μ at a 95% level of confidence. However we have no control on the width of the confidence interval. If s is large the width $\frac{2st_0}{n-1}$ might exceed 0.1. The way we proceed is that if the width is too large we draw an additional sample of size $m = m(s)$, that depends on s. We use the combined mean mean $\bar{y} = \frac{1}{n+m}\sum x_i$, but the old standard deviation s and construct the new statistic

$$T_{n-1} = \frac{\bar{y}-\mu}{s}\sqrt{n+m}\frac{\sqrt{n-1}}{\sqrt{n}}$$

It is seen to have the same t distribution with $n-1$ degrees of freedom no matter what the choice of $m = m(s)$ we make. This is because the conditional distribution of $\bar{y}$ given s is normal with mean μ and variance $\frac{1}{n+m(s)}\sigma^2$ and the scaling is done correctly. Therefore for any choice of $m(s)$ the interval

$$(\bar{y} - \frac{st_0\sqrt{n}}{\sqrt{(n-1)(n+m)}}, \bar{y} + \frac{st_0\sqrt{n}}{\sqrt{(n-1)(n+m)}})$$

is a 95% confidence interval for μ of width $2\frac{st_0\sqrt{n}}{\sqrt{(n-1)(n+m)}}$ which can be made less than 0.1 by taking $m(s)$ so that

$$\frac{2st_0\sqrt{n}}{\sqrt{(n-1)(n+m)}} \leq 0.1$$

or

$$m(s) \geq 400 s^2 t_0^2 \frac{n-1}{n} - n$$

24 Sequential Procedures

Let us imagine a situation where we want to examine products from an assembly line to be sure that the percentage of defective items is no more than 1%. Let us examine randomly selected items one by one. If the first three selected items are defective it seems hardly worthwhile to continue, whereas if the first three are good it does not really convince us. The typical procedure in this kind of a situation is to draw observations one at atime. After each observation either make a decision one way or another, or decide to continue and make one more observation.

Let us look at the problem of testing the null hypothesis that the observations are drwan from f_0 against the alternative that they are drawn from f_1, wher $f_0(x)$ and $f_1(x)$ are two probability densites on R. The procedure known as the sequential probablity ratio test (SPRT) is the following. Take two numbers A_0 and A_1 such that $A_0 < 1 < A_1$. and for any n look at the ratio

$$R_n(x_1, \ldots, x_n) = \frac{f_1(x_1) \cdots f_1(x_n)}{f_0(x_1) \cdots f_0(x_n)}$$

Inductively, starting from $n = 1$,

If	$R_n < A_0$	choose H_0
If	$R_n > A_1$	choose H_1
If	$A_0 \leq R_n \leq A_0$	continue and take one more observation

It is conceivable that the procedure may never end. Of course if $f_0 = f_1$ it never will. Assuming that $f_0 \neq f_1$ let us show that the SPRT terminates with probability one, under both H_0 and H_1

Consider $Y_i = \log \frac{f_1(X_i}{f_0(X_i)}$. Under both H_0 and H_1, these are independent and identically distributed random variables such that $P[|Y_i| \geq \ell] \geq p$ for some $\ell > 0$ and $p > 0$. Then either $P[y_i \geq \ell]$ or $P[Y_i \leq -\ell] \geq \frac{p}{2}$. There is then a positive probability of getting a run of k, Y-values such that

$$P[|Y_1 + Y_2 + \cdots + Y_k| \geq k\ell] \geq (\frac{p}{2})^k$$

Pick k such that $kl \geq A_1 - A_0$. Since such a run will occur sooner or later the sums will get out definitely by then. Sine the probability of having to wait for a long time for some thing with positive probabilty to happen decays geometrically, actually the termination time N which is a random variable has a finite expectation under H_0 as well as H_1.

The SPRT will have type I and type II errors like any other test. $\alpha_0(A_0, A_1)$ and $\alpha_1(A_0, A_1)$. It is clear they can be controlled somehow by changing A_0 and A_1, and it is safer to make A_0 small and A_1 large. But this will force $E(N)$ to be large. Thers is no free lunch!

Let us try to estimate α_0 and α_1 in terms of A_0 and A_1. We denote for each n, the sets

$$\begin{aligned}
D_n(0) &= \Big\{(x_1, x_2, \ldots, x_{n-1}) \in C_{n-1}, R_n(x_1, x_2, \ldots, x_n) < A_0\Big\} \\
D_n(1) &= \Big\{(x_1, x_2, \ldots, x_{n-1}) \in C_{n-1}, R_n(x_1, x_2, \ldots, x_n) > A_1\Big\} \\
C_n &= \Big\{(x_1, x_2, \ldots, x_{n-1}) \in C_{n-1}, A_0 \leq R_n(x_1, x_2, \ldots, x_n) \leq A_1\Big\}
\end{aligned}$$

The errors are

$$\begin{aligned}
\alpha_0(A_0, A_1) &= \sum_{n=1}^{\infty} \int_{D_n(1)} f_0(x_1) \cdots f_0(x_n) dx_1 dx_2 \cdots dx_n \\
\alpha_1(A_0, A_1) &= \sum_{n=1}^{\infty} \int_{D_n(0)} f_1(x_1) \cdots f_1(x_n) dx_1 dx_2 \cdots dx_n
\end{aligned}$$

Note that on $D_n(1)$, $R_n > A_1$ and on $D_n(0)$, $R_n < A_0$. Therefore,

$$\begin{aligned}
\alpha_0(A_0, A_1) &\leq \frac{1}{A_1} \sum_{n=1}^{\infty} \int_{D_n(1)} f_1(x_1) \cdots f_1(x_n) dx_1 dx_2 \cdots dx_n \\
&= \frac{1 - \alpha_1(A_0, A_1)}{A_1} \\
\alpha_1(A_0, A_1) &\leq A_0 \sum_{n=1}^{\infty} \int_{D_n(0)} f_0(x_1) \cdots f_0(x_n) dx_1 dx_2 \cdots dx_n \\
&= A_0(1 - \alpha_0(A_0, A_1))
\end{aligned}$$

In practice on chooses

$$A_0 = \frac{\alpha_1}{1-\alpha_0}, \qquad A_1 = \frac{1-\alpha_1}{\alpha_0}$$

Let us see how this is applied in practice. Suppose we have observations on the life of a product that we are testing. The density is

$$\frac{1}{a}e^{-\frac{x}{a}}$$

for $x \geq 0$ and a is the expected life. We would like to be sure that $a = 1$. But want to be sure that the probability of accepting a product with $a < .9$ is no more than 0.01, while if a product is good enough, i.e. $a \geq 1.1$ we want the probability of rejecting it to be no more than 0.1.

We then have

$$\begin{aligned}
f_1 &= \frac{1}{0.9}\exp[-\frac{x}{0.9}] \\
f_0 &= \frac{1}{1.1}\exp[-\frac{x}{1.1}] \\
\alpha_0 &= 0.10 \\
\alpha_1 &= 0.01 \\
A_0 &= \frac{0.01}{1-0.10} \simeq 0.01 \\
A_1 &= \frac{1-0.01}{.10} \simeq 9.9
\end{aligned}$$

$$R_n = (\frac{1.1}{0.9})^n \exp[(\frac{1}{1.1} - \frac{1}{0.9})\sum_i x_i] = (\frac{1.1}{0.9})^n \exp[-0.2\sum_i x_i\ ]$$

The region $R_n < A_0$ becomes

$$\sum_i x_i > 5n\log\frac{1.1}{0.9} + 5\log 100$$

and $R_n > A_1$ becomes

$$\sum_i x_i < 5n\log\frac{1.1}{0.9} - 5\log 9.9$$

If the first alternatve happens first we accept the product. If the second alternative occurs first we reject the product. Of course we continue testing if we neither has occured.