Lecture 33: SPRT for Bernoulli and binomial populations

Statistical Inference–II ▪ Units II–III: Sequential Analysis

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

This lecture

Previous lecture

Lecture 32: The zero-drift case.

Today

- Distribution-Specific SPRTs
  - Bernoulli (binomial) distribution
  - Binomial observations at each stage

Distribution-specific SPRTs: what is given

For each model we give the log-likelihood ratio W_m, the acceptance and rejection lines, the O.C. root in the parametric form $\theta(h)$, the drift $\mu_Z(\theta)$, the zero-drift value θ^* and $\mathbb{E}_{\theta^*}(Z_1^2)$. The O.C. and A.S.N. functions then follow from the two preceding sections:

$$Q(\theta) \approx \frac{A^h - 1}{A^h - B^h}, \qquad \mathbb{E}_\theta(N) \approx \frac{Q(\theta)\log B + \{1 - Q(\theta)\}\log A}{\mu_Z(\theta)},$$

$$\mathbb{E}_{\theta^*}(N) \approx -\frac{(\log A)(\log B)}{\mathbb{E}_{\theta^*}(Z_1^2)}.$$

Throughout, the alternative is taken to lie to the right of the null value ($p_1 > p_0$, $\lambda_1 > \lambda_0$, and so on); the other case is obtained by reversing the inequalities.

Bernoulli: likelihood ratio

Let $X_i \sim \text{Bernoulli}(p)$ and test $H_0 : p = p_0$ against $H_1 : p = p_1$, $p_1 > p_0$. Put $T_m = \sum_{i=1}^{m} X_i$. Then

$$\mathcal{L}_{j,m} = p_j^{T_m}(1-p_j)^{m-T_m}, \qquad j = 0, 1,$$

so

$$\Lambda_m = \left(\frac{p_1}{p_0}\right)^{T_m} \left(\frac{1-p_1}{1-p_0}\right)^{m-T_m},$$

. . . and

$$W_m = T_m \log \frac{p_1}{p_0} + (m - T_m) \log \frac{1 - p_1}{1 - p_0}.$$

The one-step increment is

$$Z_i = X_i \log \frac{p_1}{p_0} + (1 - X_i) \log \frac{1 - p_1}{1 - p_0}.$$

Acceptance and rejection lines. Put

$$g_1 = \log \frac{p_1}{p_0} > 0, \qquad g_2 = \log \frac{1 - p_0}{1 - p_1} > 0.$$

Bernoulli: acceptance and rejection lines

Then $W_m = (g_1 + g_2)T_m - mg_2$, and $\log B < W_m < \log A$ is equivalent to

$$-\ell_0 + s\,m < T_m < \ell_1 + s\,m,$$
$$s = \frac{g_2}{g_1 + g_2}, \quad \ell_0 = \frac{-\log B}{g_1 + g_2}, \quad \ell_1 = \frac{\log A}{g_1 + g_2}.$$

Accept H_0 as soon as $T_m \le -\ell_0 + sm$ and reject H_0 as soon as $T_m \ge \ell_1 + sm$. The slope s always lies between p_0 and p_1. (The intercepts are written ℓ_0, ℓ_1 to avoid confusion with the O.C. root h.)

O.C. function. For a true value p, the O.C. root is determined by

$$p\left(\frac{p_1}{p_0}\right)^h + (1-p)\left(\frac{1-p_1}{1-p_0}\right)^h = 1,$$

which is linear in p. Solving,

$$\boxed{p(h) = \frac{1 - \left(\dfrac{1-p_1}{1-p_0}\right)^h}{\left(\dfrac{p_1}{p_0}\right)^h - \left(\dfrac{1-p_1}{1-p_0}\right)^h}, \qquad h \neq 0.}$$

One checks that $p(1) = p_0$ and $p(-1) = p_1$.

$$\mu_Z(p) = p \log \frac{p_1}{p_0} + (1-p) \log \frac{1-p_1}{1-p_0} = p(g_1 + g_2) - g_2.$$

The zero-drift value is

$$p^* = \frac{g_2}{g_1 + g_2} = \frac{\log\{(1-p_0)/(1-p_1)\}}{\log\{p_1(1-p_0)/[p_0(1-p_1)]\}},$$

which is also the limit of $p(h)$ as $h \to 0$ and coincides with the slope s of the boundary lines. At $p = p^*$,

$$Q(p^*) \approx \frac{\log A}{\log A - \log B}.$$

Bernoulli: A.S.N. at the zero-drift value

Since $\mathbb{E}_{p^*}(Z_1) = 0$,

$$\mathbb{E}_{p^*}(Z_1^2) = \mathrm{Var}_{p^*}(Z_1) = (g_1 + g_2)^2 p^*(1 - p^*) = g_1 g_2,$$

$$\mathbb{E}_{p^*}(N) \approx -\frac{(\log A)(\log B)}{g_1 g_2}.$$

(The identity $(g_1 + g_2)^2 p^*(1 - p^*) = g_1 g_2$ follows from $p^* = g_2/(g_1 + g_2)$ and $1 - p^* = g_1/(g_1 + g_2)$.)

Binomial observations at each stage

If each stage records $X_i \sim \mathrm{Binomial}(r, p)$, where r is known, the binomial coefficients cancel from the likelihood ratio. With $T_m = \sum_{i=1}^{m} X_i$,

$$\begin{aligned} W_m &= T_m \log \frac{p_1}{p_0} + (mr - T_m) \log \frac{1-p_1}{1-p_0} \\ &= (g_1 + g_2) T_m - mr\, g_2, \end{aligned}$$

so the continuation region is $-\ell_0 + s\, rm < T_m < \ell_1 + s\, rm$, with s, ℓ_0, ℓ_1 as in the Bernoulli case. Since

$$\mathbb{E}_p(e^{hZ_1}) = \left[p \left(\frac{p_1}{p_0} \right)^h + (1-p) \left(\frac{1-p_1}{1-p_0} \right)^h \right]^r,$$

Binomial observations at each stage (cont.)

. . . and a positive number raised to the power r equals one only if the number itself equals one, the O.C. root $h(p)$ and hence the O.C. curve are the *same* as in the Bernoulli case.

The drift is multiplied by r, so the A.S.N. is approximately the Bernoulli A.S.N. divided by r: each binomial stage carries the information of r Bernoulli trials. The Bernoulli case is $r = 1$.

After this lecture

Practice (Lecture Notes, Chapter 2)

- Exercises (Section 2.13): 2.1, 2.2, 2.7, 2.16.

Next lecture

Lecture 34: SPRT for a Poisson population.