Appendix A Counter-example

Assume

$$\overline{\theta} = \gamma^+ \times \theta_{g^+} + (1 - \gamma^+) \times \theta_{g^-} \tag{A.1}$$

holds and prevalence γ^+ of the marker-positive g^+ subpopulation is 0.5%. Suppose Y^+ is a Normally distributed outcome with mean θ_{g^+} and variance one. Suppose Y_1^-, Y_2^- are i.i.d. Normally distributed outcomes with mean θ_{g^-} and variance one.

Consider testing the null hypotheses

$$\bar{H}_0: \bar{\theta} = 0, \quad H_{0+}: \theta_{g^+} = 0, \quad H_{0-}: \theta_{g^-} = 0.$$

Because of (A.1), there are only four closed testing null hypotheses

$$\overline{H}_0: \overline{\theta} = 0, \ H_{0+}: \theta_{g^+} = 0, \ H_{0-}: \theta_{g^-} = 0, \ \text{and} \ H_{00}: \theta_{g^+} = \theta_{g^-} = \overline{\theta} = 0.$$

Recognizing that a level- α test for $H_{0+}: \theta_{g^+} = 0$ is also a level- α test for $H_{00}: \theta_{g^+} = \theta_{g^-} = \overline{\theta} = 0$, consider the following rejection regions which control FWER at 5%:

$$\begin{split} \text{Reject} \bar{H}_0 : \overline{\theta} &= 0 \text{ if } \quad \frac{0.005 \times Y^+ + 0.995 \times Y^-_1}{\sqrt{0.005^2 + 0.995^2}} \quad > 1.645 \\ \text{Reject} H_{0+} : \theta_{g^+} &= 0 \text{ if } \qquad Y^+ \qquad > 1.645 \\ \text{Reject} H_{0-} : \theta_{g^-} &= 0 \text{ if } \qquad Y^-_2 \qquad > 1.645 \\ \text{Reject} H_{00} : \theta_{g^+} &= \theta_{g^-} &= \overline{\theta} &= 0 \text{ if } \qquad Y^+ \qquad > 1.645 \end{split}$$

Provided $H_{00}: \theta_{g^+} = \theta_{g^-} = \overline{\theta} = 0$ is rejected, the inferences we make are

Suppose $\theta_{g^+} = 4.0$ and $\theta_{g^-} = -0.2$, then with $\theta_{g^+} > 0$, $\theta_{g^-} < 0$, and $\overline{\theta} = 0$, inferring either $\overline{\theta} > 0$ or $\theta_{g^-} > 0$ or both constitute an Incorrect Decision. With $\theta_{g^+} = 4.0$, the probability that $H_{00}: \theta_{g^+} = \theta_{g^-} = \overline{\theta} = 0$ is rejected exceeds 0.99, and the probability of making an Incorrect Decision turns out to be about 6.4% (which is greater than 5%), illustrating that controlling the FWER for testing exact equalities may not control the probability of an Incorrect Decision.