16.4 Bayesian Hypothesis Testing

- Like any form of Bayesian inference, Bayesian hypothesis testing is based on the posterior distribution.
- Define Ω_0 to be the parameter space implied by the null H_0 and Ω_a to be the parameter space implied by H_a.
- A Bayesian test compares the posterior probability $P(\theta \in \Omega_0 \mid \underset{\sim}{y})$ to $P(\theta \in \Omega_a \mid \underset{\sim}{y})$.
- We might reject H_0 when $P[H_0 \text{ true} \mid \underset{\sim}{y}]$ is less than $P[H_a \text{ true} \mid \underset{\sim}{y}]$.
- If the cost of incorrectly rejecting H_0 is high, we may choose to reject H_0 only when

$P[H_0 \text{ true} \mid \underset{\sim}{y}]$ is "much" less than $P[H_a \text{ true} \mid \underset{\sim}{y}]$ (for example, if

<u>Example 1 again:</u> Test $H_0: p \leq 0.8$ vs. $H_a: p > 0.8$ in the plant science example. Assume a beta(3,1) prior on p.

- We derived the posterior for p to be

Example 2 again: Test $H_0: \mu \leq 4$ vs. $H_a: \mu > 4$ where μ is the mean waiting time.

- The case of a "point null" and a two-tail alternative, e.g.,

is trickier, since if $g^*(\theta)$ is continuous,

regardless of the data.

- A common approach in this case is to form a two-sided $100(1-\alpha)\%$ credible interval for θ using some reasonable α, and to reject H_0 if:

<u>Example 3:</u> Test $H_0: \mu = 90$ vs. $H_a: \mu \neq 90$ using our $N(100, 25)$ prior for μ.

<u>Note:</u> The concepts of size (i.e., P[Type I error]) and power of a test are not of concern to Bayesians, who do not consider θ as having a fixed value in repeated sampling.

16.5 Additional Comments

- We have used conjugate priors in our examples because with such priors it is easy to derive the posterior.
- The analyst may certainly use non-conjugate priors, but in that case, the posterior distribution often cannot be derived analytically and must be approximated by simulation methods.

Note: The marginal pdf (or pmf) of the data $m(y_1, \dots, y_n)$ is also called the predictive distribution of the data.

- It displays the pattern the data should follow, averaged over the possible values of the parameter θ.
- A common approach to check model fit is to compare the data actually observed, $y_1, \dots, y_n$, to the predictive distribution.

- If $m(y_1, \ldots, y_n)$ provides a poor fit to our observed $y_1, \ldots, y_n$, then perhaps either the prior or the data model should be reconsidered.
- Bayesian estimators can be evaluated using criteria with which we judged estimators in STAT 512.
- Bayesian point estimators are typically biased (but they are usually consistent).
- We know that the MSE of an estimator $\hat{\theta}$:

  can be used to compare biased and unbiased estimators.
- Recall in STAT 512, for $Y_1, \ldots, Y_n \overset{iid}{\sim}$ Bernoulli (p), we compared

  on the basis of MSE.

- Note that $\hat{p}^*$ is the Bayes estimate if we use a ____________ prior on p (i.e., a ____________ prior).
- We found the MSE of $\hat{p}^*$ was lower (better) than the MSE of $\hat{p}$ for certain values of p (roughly when
- In general, Bayes estimators will have lower MSE than classical MVUEs at least for some values of the target parameter θ.
- A common definition of the "Bayes risk" of an estimator is the expected value of the MSE with respect to the posterior distribution of θ:

- The posterior mean is the estimator $\hat{\theta}$ that minimizes this Bayes risk.

- Bayes estimators also have good large-sample properties:

- Under usual regularity conditions, as $n \to \infty$, Bayes estimators (like MLEs) are:

  * consistent
  * asymptotically normal
  * asymptotically efficient.