Hypothesis Testing - Normal Example

We have 2 batches of paint, one of which is quick-dry; quick-dry paint will dry in an average of 10 minutes, and regular paint in an average of 25 minutes. The paint is unlabeled, and we forgot which was which! We paint 5 boards from batch 1 and record the drying time. We believe that the drying times are normally distributed with a standard deviation of 5 minutes for both paint types.

We have a vague feeling that batch 1 is quick-dry; we will calculate a p-value to see how strong the evidence is that the batch we used is *not* 25 minutes:

$H_0 : \theta = 25$

$H_A : \theta = 10$

(a) Find the form of the likelihood ratio statistic for this test.

You will need to use the fact that if $X_i \sim \text{Normal}(\theta, \sigma^2)$ then its pdf is $f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-1}{2\sigma^2}(x_i - \theta)^2\right]$. In this case, we are assuming $\sigma = 5$; the value of θ will be different in the numerator and denominator of the likelihood ratio statistic.

$$W = \frac{\mathcal{L}(\theta_0 | X_1, \dots, X_5)}{\mathcal{L}(\theta_A | X_1, \dots, X_5)} = \frac{f_{X_1, \dots, X_5}(X_1, \dots, X_5 | \theta_0)}{f_{X_1, \dots, X_5}(X_1, \dots, X_5 | \theta_A)}$$

$$= \frac{\prod_{i=1}^{5} \cancel{\frac{1}{\sqrt{2\pi\sigma^2}}} \exp\left[\frac{-1}{2\sigma^2}(X_i - 25)^2\right]}{\prod_{i=1}^{5} \cancel{\frac{1}{\sqrt{2\pi\sigma^2}}} \exp\left[\frac{-1}{2\sigma^2}(X_i - 10)^2\right]}$$

(b) Show that the likelihood ratio statistic can be simplified to a form that involves the sample mean $\bar{X}$.

$$W = \frac{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5}(X_i - 25)^2\right]}{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5}(X_i - 10)^2\right]} = \frac{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5}\left\{X_i^2 - 2\cdot 25\cdot X_i + 25^2\right\}\right]}{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5}\left\{X_i^2 - 2\cdot 10\cdot X_i + 10^2\right\}\right]}$$

$$= \frac{\cancel{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5} X_i^2\right]} \cdot \exp\left[\frac{-1}{2\sigma^2}(-2\cdot 25)\sum_{i=1}^{5} X_i\right] \cdot \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5} 25^2\right]}{\cancel{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5} X_i^2\right]} \cdot \exp\left[\frac{-1}{2\sigma^2}(-2\cdot 10)\sum_{i=1}^{5} X_i\right] \cdot \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5} 10^2\right]}$$

$$= \frac{\exp\left[\frac{\cancel{-}1}{\cancel{2}\sigma^2}(\cancel{-2}\cdot 25)\cdot 5\cdot \bar{X}\right] \cdot \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5} 25^2\right]}{\exp\left[\frac{\cancel{-}1}{\cancel{2}\sigma^2}(\cancel{-2}\cdot 10)\cdot 5\bar{X}\right] \cdot \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{5} 10^2\right]}$$

(c) Show how the p-value for the likelihood ratio test could be calculated based on an observed sample mean of $\bar{x}$. Your answer should be a probability involving the random variable $\bar{X}$ and the observed value $\bar{x}$ based on the observed sample data.

given H_0 correct

$$\text{p-value} = P(W \le w \mid H_0 \text{ correct})$$

$$= P\left(\frac{\exp\left[\frac{25}{\sigma^2} \cdot 5\bar{X}\right] \cdot \cancel{\exp\left[\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{5} 25^2\right]}}{\exp\left[\frac{10}{\sigma^2} \cdot 5\bar{X}\right] \cdot \cancel{\exp\left[\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{5} 10^2\right]}} \le \frac{\exp\left[\frac{25}{\sigma^2} 5\bar{x}\right] \cdot \cancel{\exp\left[\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{5} 25^2\right]}}{\exp\left[\frac{10}{\sigma^2} \cdot 5\bar{x}\right] \cdot \cancel{\exp\left[\frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{5} 10^2\right]}} \right)$$

$$= P\left(\exp\left[\frac{5\bar{X}}{\sigma^2}(25-10)\right] \le \exp\left[\frac{5\bar{x}}{\sigma^2}(25-10)\right] \right)$$

(d) Show that your p-value calculation from part (c) is equivalent to $P(\bar{X} \le \bar{x})$

continuing above, taking a log:

$$\text{p-value} = P\left(\frac{\cancel{5}\bar{X}}{\cancel{\sigma^2}} \boxed{\cancel{\cdot 15}} \le \frac{\cancel{5}\bar{x}}{\cancel{\sigma^2}} \boxed{\cancel{\cdot 15}} \right)$$

$$= P(\bar{X} \le \bar{x}) \quad \checkmark$$

$\theta_0 - \theta_A > 0$
so cancellation didn't change direction of inequality

If we had been testing
$H_0: \theta = 10$
$H_A: \theta = 25$

$\theta_0 - \theta_A = -15 < 0$

we would have had to change direction of inequality.

$$\text{p-value} = P(\bar{X} \ge \bar{x})$$

In both cases, $\text{p-value} = P(W \le w)$

(likelihood ratio statistic)

(e) **Is a small or large value of $\bar{x}$ evidence against H_0? How is this determined by the set up of the null and alternative hypotheses?**

A small value of $\bar{x}$ is evidence against Ho:

→ a small p-value is evidence against Ho

→ a smaller value of $\bar{x}$ will give usa smaller p-value = $\underline{P(\bar{X} \le \bar{x})}$

$\theta_A = 10$ from alternative, smaller than $\theta_0 = 25$ from Ho.

A vale of $\bar{x}$ that is small will be more consistent with H_A than Ho, so would provide evidence against Ho.