Suppose $X_1, \dots, X_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$

Both μ and σ^2 unknown.

Full parameter vector is $\underset{\sim}{\Theta} = (\mu, \sigma^2)$

Parameter space is $\Omega = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 \geq 0\}$

Consider a test of

$H_0: \mu = \mu_0$

$H_A: \mu \neq \mu_0$

$H_0: \underset{\sim}{\Theta} \in \Omega_0 = \{(\mu, \sigma^2) : \mu = \mu_0, \sigma^2 \geq 0\}$

$H_A: \underset{\sim}{\Theta} \in \Omega_A = \{(\mu, \sigma^2) : \mu \neq \mu_0, \sigma^2 \geq 0\}$

Reminder of likelihood ratio statistic:

$$W = \frac{\mathcal{L}(\Theta_0 | X_1, \dots, X_n)}{\mathcal{L}(\Theta_A | X_1, \dots, X_n)}$$ relevant for simple vs simple hypotheses.

Generalized LR statistic useful for composite hypotheses:

$$W = \frac{\max_{\underset{\sim}{\Theta} \in \Omega_0} \mathcal{L}(\underset{\sim}{\Theta} | X_1, \dots, X_n)}{\max_{\underset{\sim}{\Theta} \in \Omega} \mathcal{L}(\underset{\sim}{\Theta} | X_1, \dots, X_n)}$$

$\leftarrow$ the max. of the likelihood, subject to the constraint that H_0 is true. (numerator)

$\leftarrow$ the max. of the likelihood (denominator)

For normal example,

- for denominator, we've previously shown that the max. lik. estimators are $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$
- for numerator, know $\mu = \mu_0$. Given that, can show that MLE for σ^2 is $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_0)^2$

For a particular value of $\underset{\sim}{\theta} = (\mu, \sigma^2)$, the likelihood is

$$\mathcal{L}(\underset{\sim}{\theta} \mid \underset{\sim}{X}) = \mathcal{L}(\mu, \sigma^2 \mid \underset{\sim}{X}) = f_{\underset{\sim}{X} \mid \mu, \sigma^2}(X_1, \ldots, X_n \mid \mu, \sigma^2)$$

$$= \prod_{i=1}^{n} f_{X_i \mid \mu, \sigma^2}(X_i \mid \mu, \sigma^2)$$

$$= \prod_{i=1}^{n} (2\pi)^{-1/2} (\sigma^2)^{-1/2} \cdot \exp\left\{ \frac{-1}{2\sigma^2} (X_i - \mu)^2 \right\}$$

$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 \right\}$$

The generalized likelihood ratio statistic is:

$$W = \frac{\mathcal{L}\left(\mu_0, \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2 \mid \underset{\sim}{X}\right)}{\mathcal{L}\left(\bar{X}, \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \mid \underset{\sim}{X}\right)}$$

$$= \frac{\cancel{(2\pi)^{-n/2}} \left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2\right)^{-n/2} \exp\left[\frac{-1}{\cancel{2\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2}} \cdot \cancel{\sum_{i=1}^{n}(X_i - \mu_0)^2}\right]}{\cancel{(2\pi)^{-n/2}} \left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^{-n/2} \exp\left[\frac{-1}{\cancel{2\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}} \cdot \cancel{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]}$$

$$= \left\{ \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2} \right\}^{n/2}$$

---

p-value $= P(\underline{W} \le \underline{w} \mid H_0 \text{ true})$ $\quad \mu = \mu_0$

Generalized L.R. statistic, as a r.v. based on $X_1, \ldots, X_n$

G. L. R. statistic based on observed data $x_1, \ldots, x_n$

$$\underline{p\text{-value}} = P\left(\left\{\frac{\cancel{\frac{1}{n}}\sum_{i=1}^{n}(X_i-\bar{X})^2}{\cancel{\frac{1}{n}}\sum_{i=1}^{n}(X_i-\mu_0)^2}\right\}^{n/2} \le \left\{\frac{\cancel{\frac{1}{n}}\sum_{i=1}^{n}(x_i-\bar{x})^2}{\cancel{\frac{1}{n}}\sum_{i=1}^{n}(x_i-\mu_0)^2}\right\}^{n/2} \middle| \mu=\mu_0\right)$$

$$= P\left(\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sum_{i=1}^{n}(X_i-\mu_0)^2} \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2} \middle| \mu=\mu_0\right)$$

$$= P\left(\frac{\sum_{i=1}^{n}(X_i-\mu_0)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \ge \frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \middle| \mu=\mu_0\right)$$

(note: if $\frac{1}{5} \le \frac{1}{3}$ then $5 \ge 3$)

$$= P\left(\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2 + n(\bar{X}-\mu_0)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \ge \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \middle| \mu=\mu_0\right)$$

(previous result: $\sum(x_i-\mu)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$

$$= P\left(\cancel{1+}\frac{n(\bar{X}-\mu_0)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \ge \cancel{1+}\frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \middle| \mu=\mu_0\right)$$

$$= P\left(\frac{(\bar{X}-\mu_0)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2/n} \ge \frac{(\bar{x}-\mu_0)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2/n} \middle| \mu=\mu_0\right)$$

$$= P\left(\frac{|\bar{X}-\mu_0|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}/\sqrt{n}} \ge \frac{|\bar{x}-\mu_0|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}/\sqrt{n}} \middle| \mu=\mu_0\right)$$

$$= P\left(T \le \frac{-|\bar{x}-\mu_0|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}/\sqrt{n}} \text{ or } T \ge \frac{|\bar{x}-\mu_0|}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}/\sqrt{n}} \,\middle|\, \mu = \mu_0 \right)$$

where $T = \dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$ (the t statistic)