Last time: Vocabulary.

Hypothesis: proposition that is true/false

- $H_0 = \text{null hypothesis} \subseteq \text{want to reject}$
- $H_1 = \text{alternative hypothesis} \subseteq \text{reject}$

$\alpha = P(\text{Type I}) \equiv "\text{size}" = \text{level of significance} = P(H_0 \text{ rejected but true}) = P(C \cap H_0)$

$\beta = P(\text{Type II}) = P(H_0 \text{ accepted and } H_1 \text{ false}) = P(C \cap H_1)$

$1 - \beta = \text{"power of test"}$

E.g.:

$H_0: \theta = .9$

$H_1: \theta = .6$

$c = \{ X < 14 \}$

out of $n = 20$

Def.: Simple vs Composite Hypothesis.

- Simple: $H_0$ completely determines pdf sample.

Want to test $H_0$ against $H_1$. E.g.: Want estimator for $\mu$ from $X_1, X_2$. $\bar{X} = \frac{1}{2} (X_1 + X_2)$.

$c = \{ X > K \}$

$X_1, X_2 \sim \text{Normal} \Rightarrow H_0: \mu = \mu_0$
\( z \geq 2x_1 - x_2 \geq 2K \)

Last time: chose \( K \) so that

\[ \Pr(\mathbb{C} \mid H_0) = \Pr(\mathbb{C} \cap H_0) \]

Which of these regions is most powerful for \( H_1 \)?

Recall: likelihood

\[ L(x_1, \ldots, x_n ; H_0) = L_0 = \frac{f(x_1)}{x_1 \mid \mathbb{H}_0} \cdots \frac{f(x_n)}{x_n \mid \mathbb{H}_0} \]

E.g. \( H_0: X_i \text{ are } N(\mu_0, 1) \)

\[ L_0(x_1, \ldots, x_n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i - \mu_0)^2}{2}} \]

\( H_1: X_i \text{ are } N(\mu_1, 1) \)

Neyman–Pearson Lemma: Suppose \( \mathbb{C} \) is such that:

1. \( \frac{L_0(x)}{L_1(x)} \leq \kappa \forall x \in \mathbb{C} \), and
2. \( \frac{L_0(x)}{L_1(x)} > \kappa \forall x \notin \mathbb{C} \)

Then \( \mathbb{C} \) is a most powerful critical region for \( H_0 \) vs \( H_1 \).

E.g.: Construct a most powerful critical region.
Look at \( \frac{L_0}{L_1} = e^{-\frac{(x_0-x_0)^2}{2}} \cdots e^{-\frac{(x_n-x_0)^2}{2}} \)

\[ = e^{-\frac{1}{2} \left[ x_1^2 - 2x_1m_0 + m_0^2 + x_2^2 - 2x_2m_0 + m_0^2 + \cdots + x_n^2 - 2x_nm_0 + m_0^2 \right]} \]

\[ = e^{-\frac{1}{2} \left[ x_1^2 - x_1 + m_1^2 + x_2^2 - x_2 + m_1^2 + \cdots + x_n^2 - x_n + m_1^2 \right]} \]

\[ = e^{\frac{1}{2} \left[ px_1, m_0 \right] + px_1, m_1 + px_2, (m_0, m_1) + \cdots + px_n, (m_0, m_1)} \]

\[ = e^{-\frac{1}{2} \left[ n m_0^2 - n m_1^2 \right]} \]

\[ = \left( e^{n m_1 n} \right) \left( m_0 - m_1 \right)
\]

\[ = \left( e^{n m_1 n} \right) \left( m_0 - m_1 \right) \cdot n \cdot x \]

\[ = \left[ \left( m_0 - m_1 \right) \cdot n \cdot x \right] + \text{constant} \]

So \( \frac{L_0}{L_1} \) is a function of \( x \) so it \( \frac{L_0}{L_1} \) \( x \) inside \( \frac{L_0}{L_1} \) \( x \) inside

\[ \text{then } C: \{ x > k \} \]

\[ \& \frac{L_0}{L_1} \text{ outside} \]

\[ \text{outside} \]
Proof of Lemma: Suppose $C$ is such that
\[ \frac{L_0}{L_1} \leq k \text{ in } C \Rightarrow \frac{L_0}{L_1} > k \text{ in } C^c. \] Let $D$ be another

\[ \alpha = \int_C L_0(x) \, dx = \int_D L_0(x) \, dx \]

\[ = \int_C L_0 + L_0 = \int_D L_0 + L_0 \]

\[ = \int_{C \cap D} + \int_{C \cap D} = \int_{D \cap C} + \int_{D \cap C} \]

\[ \Rightarrow \int_C L_0 = \int_D L_0 \leq k \int_D L_1 \]

\[ \Rightarrow k \int_C L_1 \leq k \int_D L_1 + \int_D L_1 \]

\[ \Rightarrow \int_{C \cap D} + \int_{D \cap C} \leq \int_{C \cap D} + \int_{D \cap C} \]

\[ \Rightarrow \int_D L_1 \leq \int_D L_1 \]
To finish the proof, note that the power of the test when using critical region $C$ is:

$$1 - \beta = 1 - \Pr(\text{Type II error})$$

$$= 1 - \Pr(C^c \mid H_1) = \Pr(C \mid H_1)$$

$$= \int_C L_1(x) \, dx$$

So the last equality,

$$\int_D L_1(x) \, dx \leq \int_C L_1(x) \, dx,$$

means that the power of $C$ is greater than that of $D$. Since $D$ was any critical region of size $\alpha$, this implies that $C$ is indeed a most powerful critical region for testing $H_0$ against $H_1$. (I was trying to say this in terms of $\beta$, but it is just as well to prove the inequality with $1 - \beta$...).