Math 373 Lecture 17
Statistical Tests of Hypothesis, p-values

Much of statistics consists of choosing between two hypotheses. One hypothesis is the null hypothesis $H_0$. It is usually the default assumption. The negation of the null hypothesis is the alternative hypothesis $H_a$. We reject the null hypothesis only when we have significant evidence to disprove the it. We stick with $H_0$ until its probability (its $p$-value) falls below the given level of significance.

To apply a statistical test, restate the null hypothesis $H_0$ as an assertion $\mu = \mu_0, \mu \leq \mu_1, \mu \geq \mu_0$, that compares some parameter $\mu$ (typically a mean, proportion, or $z$-score) with some claimed or expected value null value $\mu_0$. The alternative hypothesis $H_a$ will be the negation: $\mu \neq \mu_0, \mu < \mu_0, \mu > \mu_0$. These are the two-tailed, left-tailed and right-tailed tests respectively. The last two are called one-sided tests. Tests are named for the direction of the tail of the alternate hypothesis. $H_0$ must always be $\leq, =, or \geq$.

Even when we have an sample estimate $\bar{x}$ of $\mu$, we assume $\mu$ is the null-hypothesis value $\mu_0$.

Two-tailed test of hypothesis. The alternative hypothesis $H_a: \mu \neq \mu_0$ simply asserts that the test statistic differs from the expected $\mu_0$. In this case we accept $H_a$ and reject the null hypothesis $\mu = \mu_0$ when $\mu$ is significantly far above or below $\mu_0$ in either the right or left tail. E.g. $\mu \neq \mu_0 \pm \sigma_{x}\sigma_{Z}$.

$H_a$ is $\mu \neq \mu_0, \mu > \mu_0$, null region: $\mu \in \{\mu_0\}$. Right-tailed test. If the alternative hypothesis is $H_a: \mu > \mu_0$, we accept $H_a$ when $\mu$ is significantly far above $\mu_0$ in the right tail. E.g. $\mu > \mu_0 + \sigma_{x}\sigma_{Z}$.

$H_a$ is $\mu > \mu_0, \mu < \mu_0$, null region: $\mu \in (-\infty, \mu_0]$. Left-tailed test. Likewise, $\mu < \mu_0$ is accepted when $\mu$ is significantly far below $\mu_0$.

$H_a$ is $\mu < \mu_0, \mu \geq \mu_0$, null region: $\mu \in [\mu_0, \infty]$. Definition. The $p$-value of $\mu$ is the probability of being as far from the null region as $\mu$. It is the probability of $H_0$ given the evidence.

In the right-tailed case, the $p$-value of $\mu$ is $P(\mu < x)$. In the left-tailed case, the $p$-value of $\mu$ is $P(\mu > x)$ where $\mu$ is on the opposite side of $\mu_0$ from $\mu$.

When the $p$-value is low the null hypothesis is unlikely. The significance level $\alpha$, determines just how unlikely the null hypothesis must be to be rejected. Typically $\alpha$ will be $5\%$ or $1\%$. We reject $H_0$ iff its probability is less than the significance level $\alpha$, the $p$-value is $< \alpha$. We accept $H_a$ iff we reject $H_0$.

Suppose the null value is $\mu_0 = 50$, suppose $SE = 2$, suppose $\alpha = 5\%$ and suppose the sample mean is $\bar{x} = 48$.

(a) Find $p$-value of the null hypothesis that $\mu = 50$.

$H_a: \mu \neq 50 \text{ } H_0: \mu = 50 \text{ } \text{null region: } \mu \in [50]$

$z$-score $48 = (\bar{x} - \mu_0) / SE = (48 - 50) / 2 = -1$.

The point on the opposite side of $50$ from $48$ is $52$. The $z$-score of $52 = 1$.

$p$-value of $48 = \text{the probability of being as far from null region } \{\mu \} \text{ as } 48 = P(x \leq 48) + P(52 < x) = P(\mu < 1) + P(1 < \mu) = 2P(\mu < 1) = 2(0.1587) = .3174 = 32\%$

(b) Should we reject this null hypothesis at the $5\%$ significance level? Null interval for $p$-values: $[.05, 1]$.

No, the null hypothesis has probability ($p$-value) $32\%$ it isn’t sufficiently unlikely (less than $5\%$) to be rejected. Accept $H_a$.

(c) Find $p$-value of the null hypothesis that $\mu \geq 50$.

$H_a: \mu < 50, H_0: \mu \geq 50 \text{ } \mu_0 = 50, \text{ } \text{null region } \mu \in [50, \infty)$. This is a left-tailed test (the alternate region has a left tail).

As before $z$-score of $48 = -1$.

$p$-value of $48 = \text{the probability of being as far from the null } [50, \infty) \text{ as } 48 = P(x \leq 48) + P(\mu > 50) = 1.1587 = 16\%$.

(d) Should we reject this null hypothesis at the $5\%$ significance level? Null interval for $p$-values: $[.05, 1]$.

No, $16\%$ is not less than the required $5\%$ level for rejection. Accept $H_0$.

A lumber company claims that its $6 \times 4$’s are $6’$ long on average. Suppose a sample of $2 \times 4$’s averages $\bar{x} = 5.96$ with $SE = .03$. Is the difference between the observed and claimed value significant at the $5\%$ level? Should we reject the company’s claim?

Two-tailed test: there is no directionality, we don’t want $2 \times 4$’s which are too long or too short. Alternate $H_a: \mu \neq 6$. Null hypoth. $H_0: \mu = 6$. Null: $\mu \in [6]$

We accept the company’s claim $H_0$ unless there is significant evidence against the claim.

$z$-score of $5.96: (5.96 - 6)/.03 = -1.33$.

$p$-value: probability of being as far from the null region $\{6\}$ as $5.96 = P(x < 5.96) + P(x > 6.04) = \text{P}(x < -1.33) + \text{P}(1.33 < z) = 2 \times .0918 = .18$.

Simplification, just calculate $P(x < 5.96)$ and double it.

We don’t have sufficient evidence to reject the company’s claim. Since random chance can account for the observed sample average a substantial $18\%$ of the time, this is not significantly unusual. The $p$-value is not below the required $5\%$ significance level.

An engine manufacturer claims its engines produce at least $100$ hp. A sample average is $\bar{x} = 98$ hp with $SE = 1$. Should we reject the companies claim at the $5\%$ significance level?

Left-sided test: we reject the claim only if the observed horse power is significantly $< \text{ than claimed}$.

Alternate $H_a: \mu < 100$. Null hypothesis $H_0: \mu \geq 100$.

$z$-score of $98: -2$. Null region: $\mu \in [100, \infty)$

$p$-value: $P(x < 98) = P(z < -2) = .02$ $p$-interval: $p \in [.05, 1]$. Reject the null hypothesis. The $2\%$ $p$-value is significant at the $\alpha = 5\%$ level. $H_0$ is too unlikely to accept.