

Testing of Hypotheses

UGC NET Statistics | Unit: Statistical Inference | Concept + Formulas + Solved Examples

1. Definitions

NULL AND ALTERNATIVE HYPOTHESES

The **null hypothesis** H_0 is a statement of "no effect" or the status quo, tested against an **alternative hypothesis** H_1 (or H_a). H_1 can be one-sided ($\theta > \theta_0$ or $\theta < \theta_0$) or two-sided ($\theta \neq \theta_0$).

TYPE I AND TYPE II ERRORS

	H_0 True	H_0 False
Reject H_0	Type I Error (α)	Correct Decision
Fail to Reject H_0	Correct Decision	Type II Error (β)

$$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ true}), \quad \beta = P(\text{Fail to reject } H_0 \mid H_0 \text{ false})$$

POWER OF A TEST

$$\text{Power} = 1 - \beta = P(\text{Reject } H_0 \mid H_1 \text{ true})$$

2. Neyman-Pearson Lemma

THEOREM — NEYMAN-PEARSON LEMMA

For testing simple $H_0 : \theta = \theta_0$ vs simple $H_1 : \theta = \theta_1$, the **most powerful test** at level α rejects H_0 when the likelihood ratio is small:

$$\Lambda = \frac{L(\theta_0)}{L(\theta_1)} \leq k$$

for a constant k chosen so that $P(\Lambda \leq k \mid H_0) = \alpha$. This gives the **Uniformly Most Powerful (UMP)** test for simple-vs-simple hypotheses.

3. Common Test Statistics

Z-TEST (VARIANCE KNOWN, OR LARGE N)

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1) \text{ under } H_0$$

T-TEST (VARIANCE UNKNOWN, SMALL N)

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1} \text{ under } H_0$$

Two-sample t -test (equal variances, pooled): $T = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{1/n_1 + 1/n_2}} \sim t_{n_1+n_2-2}$

CHI-SQUARE TEST (GOODNESS OF FIT / INDEPENDENCE)

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \sim \chi_{k-1-m}^2$$

where O_i = observed frequency, E_i = expected frequency, m = number of parameters estimated.

F-TEST (COMPARING TWO VARIANCES)

$$F = \frac{S_1^2}{S_2^2} \sim F_{n_1-1, n_2-1} \text{ under } H_0 : \sigma_1^2 = \sigma_2^2$$

4. Key Points to Remember

- α and β trade off — decreasing α (stricter rejection) typically increases β (unless sample size increases too).
 - The p-value is the probability, under H_0 , of observing a test statistic at least as extreme as the one obtained. Reject H_0 if p-value $< \alpha$.
 - Failing to reject H_0 does **not** prove H_0 is true — it only means insufficient evidence to reject it.
 - A test is **UMP** (uniformly most powerful) if it has the highest power among all level- α tests, for every value of the parameter under H_1 — UMP tests don't always exist for composite alternatives.
 - Increasing sample size n decreases both α and β simultaneously (for fixed effect size) — the standard way to improve a test overall.
 - One-tailed tests are more powerful than two-tailed tests for detecting an effect in the specified direction, but they cannot detect effects in the opposite direction.
-

5. Worked Examples

EXAMPLE 1 — Z-TEST

A population has $\sigma = 15$. A sample of $n = 25$ gives $\bar{x} = 105$. Test $H_0 : \mu = 100$ vs $H_1 : \mu \neq 100$ at $\alpha = 0.05$.

Solution: $Z = \frac{105 - 100}{15/\sqrt{25}} = \frac{5}{3} \approx 1.67$. Critical value for two-tailed test at $\alpha = 0.05$:

$Z_{0.025} = 1.96$. Since $|1.67| < 1.96$, **fail to reject H_0** .

EXAMPLE 2 — T-TEST

A sample of $n = 10$ gives $\bar{x} = 52$, $s = 6$. Test $H_0 : \mu = 50$ vs $H_1 : \mu > 50$ at $\alpha = 0.05$ (using $t_{9,0.05} = 1.833$).

Solution: $T = \frac{52 - 50}{6/\sqrt{10}} = \frac{2}{1.897} \approx 1.054$. Since $1.054 < 1.833$, **fail to reject H_0** — insufficient evidence that $\mu > 50$.

EXAMPLE 3 — CHI-SQUARE GOODNESS OF FIT

A die is rolled 60 times with observed frequencies for faces 1–6: 8, 12, 9, 11, 7, 13. Test if the die is fair at $\alpha = 0.05$ (expected $E_i = 10$ each, $\chi^2_{5,0.05} = 11.07$).

Solution:

$$\begin{aligned}\chi^2 &= \frac{(8-10)^2}{10} + \frac{(12-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(7-10)^2}{10} + \frac{(13-10)^2}{10} \\ &= \frac{4 + 4 + 1 + 1 + 9 + 9}{10} = \frac{28}{10} = 2.8\end{aligned}$$

Since $2.8 < 11.07$, **fail to reject** H_0 — no evidence the die is unfair.

EXAMPLE 4 — APPLYING NEYMAN-PEARSON

$X \sim N(\theta, 1)$, single observation. Test $H_0 : \theta = 0$ vs $H_1 : \theta = 1$. Find the form of the most powerful test at level α .

Solution: Likelihood ratio:

$$\Lambda = \frac{e^{-x^2/2}}{e^{-(x-1)^2/2}} = e^{-x+1/2}$$

$\Lambda \leq k \iff -x + 1/2 \leq \log k \iff x \geq c$ for some constant c . So the most powerful test rejects H_0 when $X \geq c$, with c chosen so $P(X \geq c \mid \theta = 0) = \alpha$ — this matches the natural one-sided Z -test structure.

6. Common Pitfalls (Exam Traps)

WATCH OUT

- "Accept H_0 " is a misleading phrase — statisticians say **"fail to reject"** H_0 , since absence of evidence isn't evidence of absence.
- Use Z -test when σ is known (or n is large); use t -test when σ is unknown and n is small — mixing these gives wrong critical values.
- The p-value is **not** the probability that H_0 is true — it's the probability of the observed (or more extreme) data, assuming H_0 is true.
- Degrees of freedom for chi-square goodness of fit is $k - 1 - m$, where m parameters were estimated from the data — forgetting to subtract m is a common error.

- UMP tests exist for simple-vs-simple (Neyman-Pearson) and some one-sided composite cases, but generally **do not exist** for two-sided composite alternatives.

7. Quick Revision Summary

Concept	Key Fact
Type I error	$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ true})$
Type II error	$\beta = P(\text{Fail to reject } H_0 \mid H_0 \text{ false})$
Power	$1 - \beta$
Neyman-Pearson	Reject H_0 when likelihood ratio $\Lambda \leq k$
Z -test	σ known / large n
t -test	σ unknown, small n , $\text{df} = n - 1$
Chi-square GOF	$\sum (O_i - E_i)^2 / E_i$, $\text{df} = k - 1 - m$
F -test	Compares two variances, $F = S_1^2 / S_2^2$

[UGC NET](#)[Statistics](#)[Hypothesis Testing](#)[Neyman-Pearson](#)