Uniformly Most Powerful Tests

Statistics · Module 5 — Hypothesis Testing

Node 5.3

Position in Knowledge Graph

System: Statistics

Structural Domain: Hypothesis Testing

Node: 5.3

This node extends the Neyman–Pearson optimality result from simple to composite hypotheses. When the parametric family possesses a monotone likelihood ratio, a single test can be simultaneously most powerful against every point alternative on one side, yielding a uniformly most powerful (UMP) test. The key structural result is the Karlin–Rubin theorem, which reduces composite one-sided testing to the simple-vs-simple NP framework via the MLR property. This node requires Node 5.2 (NP Lemma) and the exponential family theory of Node 2.6.

Formal Definition

UMP Test

Consider a composite testing problem:

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1.$$

A level α test ϕ^* is **uniformly most powerful (UMP)** if

$$\mathbb{E}_\theta[\phi^*(X)] \geq \mathbb{E}_\theta[\phi(X)] \quad \forall \theta \in \Theta_1$$

for every level α test ϕ.

A UMP test is simultaneously most powerful against every point in Θ_1.

Monotone Likelihood Ratio

A family $\{P_\theta : \theta \in \Theta \subseteq \mathbb{R}\}$ with densities $\{p_\theta\}$ has **monotone likelihood ratio (MLR)** in a statistic $T(X)$ if for every $\theta_1 < \theta_2$, the ratio

$$\frac{p_{\theta_2}(x)}{p_{\theta_1}(x)}$$

is a non-decreasing function of $T(x)$ (on the set where $p_{\theta_1}(x) > 0$).

Intuition

The NP Lemma gives the best test for a single point alternative. The challenge with composite alternatives is that different point alternatives might require different critical regions.

The MLR property resolves this conflict: it guarantees that the likelihood ratio between any two parameter values is monotone in the same statistic T. Consequently, the NP test for every pair (θ_0, θ_1) with $\theta_1 > \theta_0$ takes the same form — reject when T is large. A single threshold test on T is therefore optimal against the entire composite alternative simultaneously.

The MLR condition is a coherence requirement: as θ increases, the distribution of T shifts in a consistent direction.

Structural Role

UMP tests occupy a narrow but important niche: they exist for one-sided hypotheses in MLR families, which includes all one-parameter exponential families with increasing natural parameter.

For two-sided alternatives ($H_1 : \theta \neq \theta_0$), UMP tests generally do not exist. This non-existence motivates the shift to: - Unbiased UMP (UMPU) tests for two-sided problems in exponential families. - Likelihood ratio tests (Node 5.4) as a general-purpose fallback.

The Karlin–Rubin theorem connects the exponential family structure (Node 2.6) to the testing framework (Node 5.1) through the NP optimality criterion (Node 5.2).

Mathematical Development

Theorem – Karlin–Rubin

Suppose the family $\{P_\theta : \theta \in \Theta \subseteq \mathbb{R}\}$ has MLR in $T(X)$. For testing

$$H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0,$$

the test

$$\phi^*(x) = \begin{cases} 1 & T(x) > c, \\ \gamma & T(x) = c, \\ 0 & T(x) < c, \end{cases}$$

where c and γ are chosen so that $\mathbb{E}_{\theta_0}[\phi^*(X)] = \alpha$, is UMP level α.

Proof

Step 1: Most powerful at each $\theta_1 > \theta_0$.

Fix any $\theta_1 > \theta_0$. By the Neyman–Pearson Lemma, the most powerful level α test of $H_0' : \theta = \theta_0$ vs $H_1' : \theta = \theta_1$ rejects for large values of

$$\frac{p_{\theta_1}(x)}{p_{\theta_0}(x)}.$$

By the MLR property, this ratio is non-decreasing in $T(x)$. Therefore the NP critical region $\{p_{\theta_1}/p_{\theta_0} > k\}$ is equivalent to $\{T(x) > c'\}$ for some threshold c'.

The level constraint $\mathbb{E}_{\theta_0}[\phi] = \alpha$ uniquely determines the threshold (and randomization constant) when the distribution of T under θ_0 is specified. Since this threshold depends only on θ_0 and α, not on θ_1, the same test ϕ^* is most powerful against every $\theta_1 > \theta_0$.

Step 2: Level α over all of Θ_0.

We must verify $\mathbb{E}_\theta[\phi^*(X)] \le \alpha$ for all $\theta \le \theta_0$, not just at θ_0.

The power function $\beta(\theta) = \mathbb{E}_\theta[\phi^*(X)]$ is non-decreasing (proved below). Since $\beta(\theta_0) = \alpha$, we have $\beta(\theta) \le \alpha$ for all $\theta \le \theta_0$. □

Proposition: Monotonicity of the Power Function

If $\{P_\theta\}$ has MLR in T and ϕ^* is a threshold test on T, then $\beta(\theta) = \mathbb{E}_\theta[\phi^*(X)]$ is non-decreasing in θ.

Proof. Let $\theta_1 < \theta_2$. The test ϕ^* is the NP test of P_{θ_1} vs P_{θ_2} at some level $\beta(\theta_1)$. By the NP corollary (the NP test is unbiased), $\beta(\theta_2) \ge \beta(\theta_1)$. □

MLR in Exponential Families

For a one-parameter exponential family with canonical density

$$p_\theta(x) = h(x)\exp\{\eta(\theta)T(x) - A(\theta)\},$$

if $\eta(\theta)$ is strictly increasing in θ, then for $\theta_1 < \theta_2$:

$$\frac{p_{\theta_2}(x)}{p_{\theta_1}(x)} = \exp\{[\eta(\theta_2) - \eta(\theta_1)]T(x) - [A(\theta_2) - A(\theta_1)]\},$$

which is strictly increasing in $T(x)$ since $\eta(\theta_2) - \eta(\theta_1) > 0$. Thus exponential families with increasing natural parameter have MLR in the natural sufficient statistic T.

Non-Existence of UMP for Two-Sided Alternatives

Proposition. For a one-parameter exponential family, there is no UMP level α test of $H_0 : \theta = \theta_0$ vs $H_1 : \theta \ne \theta_0$ (when $0 < \alpha < 1$).

Proof sketch. The UMP test of $\theta = \theta_0$ vs $\theta > \theta_0$ rejects for large T. The UMP test of $\theta = \theta_0$ vs $\theta < \theta_0$ rejects for small T. These are different tests, so no single test is simultaneously most powerful against alternatives on both sides. □

Unbiased UMP (UMPU) Tests

For two-sided problems in exponential families, one restricts to unbiased tests and seeks the UMP within that class. A level α test is **unbiased** if $\beta(\theta) \ge \alpha$ for all $\theta \in \Theta_1$.

For testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta \ne \theta_0$ in a one-parameter exponential family, the UMPU test rejects when $T \le c_1$ or $T \ge c_2$, with randomization at the boundaries, where c_1, c_2 are determined by

$$\mathbb{E}_{\theta_0}[\phi(X)] = \alpha, \qquad \mathbb{E}_{\theta_0}[T(X)\phi(X)] = \alpha\,\mathbb{E}_{\theta_0}[T(X)].$$

Worked Examples

Example 1: Normal Mean (Known Variance)

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$, σ^2 known. Test $H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0$.

This is a one-parameter exponential family with $T = \sum X_i$ and natural parameter $\eta = \mu/\sigma^2$, which is increasing in μ. By the Karlin–Rubin theorem, the UMP level α test rejects when $\bar{X} > c$.

Under μ_0: $\bar{X} \sim N(\mu_0, \sigma^2/n)$, so $c = \mu_0 + z_\alpha \sigma/\sqrt{n}$.

Power function:

$$\beta(\mu) = 1 - \Phi\left(z_\alpha - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right).$$

This is strictly increasing in μ, confirming monotonicity.

Example 2: Bernoulli Proportion

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$. Test $H_0 : p \leq 1/2$ vs $H_1 : p > 1/2$ at level $\alpha = 0.05$, with $n = 20$.

The sufficient statistic is $T = \sum X_i \sim \text{Binomial}(20, p)$. The Bernoulli family has MLR in T since $\eta(p) = \log(p/(1-p))$ is strictly increasing.

Under $p = 1/2$: - $P_{1/2}(T \geq 15) = \sum_{k=15}^{20} \binom{20}{k}(1/2)^{20} \approx 0.0207$. - $P_{1/2}(T \geq 14) \approx 0.0577$.

Non-randomized: reject when $T \geq 15$, with size $0.0207 < 0.05$.

Randomized UMP test: $\phi(x) = 1$ for $T \geq 15$, $\phi(x) = \gamma$ for $T = 14$, where

$$\gamma = \frac{0.05 - 0.0207}{P_{1/2}(T = 14)} = \frac{0.0293}{0.0370} \approx 0.792.$$

Power at $p = 0.7$: $\beta(0.7) = P_{0.7}(T \geq 15) + 0.792 \cdot P_{0.7}(T = 14) \approx 0.584 + 0.792 \cdot 0.192 \approx 0.736$.

Cross-Links

- **AI:** Monotone decision boundaries in UMP tests correspond to threshold classifiers on a score function. In calibrated binary classifiers under log-loss, the optimal decision rule is monotone in the predicted probability.
- **Economics:** One-sided tests arise in policy evaluation (e.g., testing whether a program has a positive effect). The UMP property ensures no other level-α test has higher power to detect the posited direction of effect.
- **Linear Algebra:** The sufficient statistic T is a linear functional of the data in exponential families; the UMP critical region is a half-space in the sufficient statistic space.
- **Quantum:** For quantum state families with a monotone structure (e.g., thermal states parameterized by temperature), the optimal measurement for one-sided discrimination is the quantum analogue of the UMP test.

Structural Compression

Invariant: MLR reduces the composite testing problem to a sequence of simple-vs-simple problems all solved by the same critical region.

Mechanism: The statistic T orders the family monotonically; the threshold test on T is simultaneously Neyman–Pearson optimal against every point alternative in H_1, because the NP critical region is the same for every $\theta_1 \in \Theta_1$.

Cross-System Echo: The monotone likelihood ratio is a stochastic ordering condition: T under P_{θ_2} is stochastically larger than under P_{θ_1} when $\theta_2 > \theta_1$. This ordering principle recurs in reliability theory (increasing failure rate families), decision theory (monotone decision procedures), and auction theory (revenue monotonicity under regular distributions).

Exercises

1. Show that the Bernoulli(θ) family has MLR in $T = \sum_{i=1}^{n} X_i$, and derive the UMP test for $H_0 : \theta \leq 1/2$ vs $H_1 : \theta > 1/2$ at level $\alpha = 0.10$ with $n = 10$.
2. Prove that the power function of the Karlin–Rubin test is non-decreasing in θ, using only the NP Lemma and the MLR property.
3. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$. Verify the MLR property in $T = \sum X_i$ and find the UMP test of $H_0 : \lambda \leq 1$ vs $H_1 : \lambda > 1$ for $n = 10$ and $\alpha = 0.05$.
4. Prove that for a one-parameter exponential family with strictly increasing $\eta(\theta)$, there is no UMP test of $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$.
5. For $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Gamma}(\alpha_0, \beta)$ with α_0 known, show that the family has MLR in $\sum X_i$ and state the UMP test for $H_0 : \beta \leq \beta_0$ vs $H_1 : \beta > \beta_0$.
6. Construct the UMPU test of $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$ for $X \sim N(\mu, 1)$ (single observation). Verify that it coincides with the two-sided z-test.
7. Argue, structurally, why the MLR property is the precise condition under which the NP lemma's pointwise optimality lifts to uniform optimality, and explain what breaks for families that fail to satisfy MLR (connecting to the non-existence result for two-sided alternatives).

Statistics · Module 5 — Hypothesis Testing · Node 5.3
Concepts: hypothesis-testing, exponential-families
Prerequisites: 5.2, 2.6
Enables: 5.4

hbar.university — own.your.cognition