Lecture 10

18.3. **Two-sided hypothesis.**

Definition 31. If $H : \Theta \geq \theta_2$ or $\Theta \leq \theta_1$ and $A : \theta_1 < \Theta < \theta_2$, then the hypothesis is two-sided. If $H : \theta_1 \leq \Theta \leq \theta_2$ and $A : \Theta > \theta_2$ or $\Theta < \theta_1$, then the alternative is two-sided.

Let us consider two-sided hypothesis.

Theorem 26 (c.f. Schervish, Thm 4.82, p. 249)**.** *In a one-parameter exponential family with natural parameter Θ, if $\Omega_H = (-\infty, \theta_1] \cup [\theta_2, \infty)$ and $\Omega_A = (\theta_1, \theta_2)$, with $\theta_1 < \theta_2$ a test of the form*

$$\phi_0(x) = \begin{cases} 1, & c_1 < x < c_2, \\ \gamma_i, & x = c_i, \\ 0, & c_1 > x \text{ or } c_2 < x, \end{cases}$$

with $c_1 \leq c_2$ minimizes $\beta_\phi(\theta)$ for all $\theta < \theta_1$ and for all $\theta > \theta_2$, and it maximizes $\beta_\phi(\theta)$ for all $\theta \in (\theta_1, \theta_2)$ subject to $\beta_\phi(\theta_i) = \alpha_i$ for $i = 1, 2$ where $\alpha_i = \beta_{\phi_0}(\theta_i)$. If $c_1, c_2, \gamma_1, \gamma_2$ are chosen so that $\alpha_1 = \alpha_2 = \alpha$, then ϕ_0 is UMP level α.

Lemma 5. *Let ν be a measure and $p_0, p_1, \dots, p_n$ ν-integrable functions. Put*

$$\phi_0(x) = \begin{cases} 1, & p_0(x) > \sum_{i=1}^n k_i p_i(x), \\ \gamma(x), & p_0(x) = \sum_{i=1}^n k_i p_i(x), \\ 0, & p_0(x) < \sum_{i=1}^n k_i p_i(x), \end{cases}$$

where $0 \leq \gamma(x) \leq 1$ and k_i are constants. Then ϕ_0 minimizes $\int [1-\phi(x)] p_0(x)\nu(dx)$ subject to the constraints

$$\int \phi(x) p_j(x) \nu(dx) \leq \int \phi_0(x) p_j(x)\nu(dx), \text{ for } j \text{ such that } k_j > 0,$$
$$\int \phi(x) p_j(x) \nu(dx) \geq \int \phi_0(x) p_j(x)\nu(dx), \text{ for } j \text{ such that } k_j < 0$$

Similarly

$$\tilde{\phi}_0(x) = \begin{cases} 0, & p_0(x) > \sum_{i=1}^n k_i p_i(x), \\ \gamma(x), & p_0(x) = \sum_{i=1}^n k_i p_i(x), \\ 1, & p_0(x) < \sum_{i=1}^n k_i p_i(x), \end{cases}$$

Then maximizes $\int [1-\phi(x)] p_0(x)\nu(dx)$ subject to the constraints

$$\int \phi(x) p_j(x) \nu(dx) \geq \int \tilde{\phi}_0(x) p_j(x)\nu(dx), \text{ for } j \text{ such that } k_j > 0,$$
$$\int \phi(x) p_j(x) \nu(dx) \leq \int \tilde{\phi}_0(x) p_j(x)\nu(dx), \text{ for } j \text{ such that } k_j < 0$$

Proof. Use Lagrange multipliers. See Schervish pp. 246-247. □

Proof of Theorem. A one parameter exponential family has density $f_{X|\Theta}(x \mid \theta) = h(x)c(\theta)e^{\theta x}$ with respect to some measure ν. Suppose we include $h(x)$ in ν (that is, we define a new measure ν' with density $h(x)$ with respect to ν) so that the density is $c(\theta)e^{\theta x}$ with respect to ν'. Then we abuse notation and write ν for ν'.

Let θ_1 and θ_2 be as in the statement of the theorem and let θ_0 be another parameter value. Define $p_i(x) = c(\theta_i)e^{\theta_i x}$ $i = 0, 1, 2$.

Suppose $\theta_0 \in (\theta_1, \theta_2)$. On this region we want to maximize $\beta_\phi(\theta_0)$ subject to $\beta_\phi(\theta_i) = \beta_{\phi_0}(\theta_i)$. Note that $\beta_\phi(\theta_i) = \int \phi(x)p_i(x)\nu(dx)$ and maximizing $\beta_\phi(\theta_0)$ is equivalent to minimizing $\int[1 - \phi(x)]p_0(x)\nu(dx)$. It seems we want to apply the Lemma with $k_1 > 0$ and $k_2 > 0$. Applying the Lemma gives the test maximizing $\beta_\phi(\theta_0)$ as

$$\phi(x) = \begin{cases} 1, & p_0(x) > \sum_{i=1}^2 k_i p_i(x), \\ \gamma(x), & p_0(x) = \sum_{i=1}^2 k_i p_i(x), \\ 0, & p_0(x) < \sum_{i=1}^2 k_i p_i(x), \end{cases}$$

Note that

$$p_0(x) > \sum_{i=1}^2 k_i p_i(x) \iff 1 > k_1 \frac{c(\theta_1)}{c(\theta_0)} e^{(\theta_1-\theta_0)x} + k_2 \frac{c(\theta_2)}{c(\theta_0)} e^{(\theta_2-\theta_0)x}.$$

Put $b_i = \theta_i - \theta_0$ and $a_i = k_i c(\theta_i)/c(\theta_0)$, and we get

$$1 > a_1 e^{b_1 x} + a_2 e^{b_2 x}.$$

We want the break points to be c_1 and c_2 so we need to solve two equations

$$a_1 e^{b_1 c_1} + a_2 e^{b_2 c_1} = 1,$$
$$a_1 e^{b_1 c_2} + a_2 e^{b_2 c_2} = 1,$$

for a_1, a_2. The solution exists (check yourself) and has $a_1 > 0$, $a_2 > 0$ as required (recall that we wanted $k_1, k_2 > 0$). So putting $k_i = a_i c(\theta_0)/c(\theta_i)$ gives the right choice of k_i in the minimizing test. Since the minimizing θ does not depend on θ_0 we get the same test for all $\theta_0 \in (\theta_1, \theta_2)$.

For $\theta_0 < \theta_1$ or $\theta_0 > \theta_2$ we want to minimize $\beta_\phi(\theta_0)$. This is done in a similar way using the second part of the Lemma.

Some work also remains to show that one can choose $c_1, c_2, \gamma_1, \gamma_2$ so that the test has level α. We omitt the details. Full details are in the proof of Theorem 4.82, p. 249 in Schervish "Theory of Statistics". □

Interval hypothesis. In this section we consider hypothesis of the form $H : \Theta \in [\theta_1, \theta_2]$ versus $A : \Theta \notin [\theta_1, \theta_2]$, $\theta_1 < \theta_2$. This will be called an interval hypothesis. Unfortunately there is not always UMP tests for testing H vs A. For an example in the case of *point hypothesis* see Example 8.3.19 in Casella & Berger (p. 392). On the other hand, comparing with the situation when the hypothesis and alternative are interchanged, one could guess that the test $\psi = 1 - \phi_0$, with ϕ_0 as in Theorem 26 is a good tests. One can show that this test satisfies a weaker criteria than UMP.

Definition 32. A test ϕ is *unbiased level* α if if has level α and if $\beta_\phi(\theta) \geq \alpha$ for all $\theta \in \Omega_A$. If ϕ is UMP among all unbiased tests it is called UMPU (uniformly most powerful unbiased) level α.

If $\Omega \subset \mathbb{R}^k$, a test ϕ is called *α-similar* if $\beta_\phi(\theta) = \alpha$ for each $\theta \in \overline{\Omega}_H \cap \overline{\Omega}_A$.

Proposition 4. *The following holds:*

(i) *If ϕ is unbiased level α and β_ϕ is continuous, then ϕ is α-similar.*
(ii) *If ϕ is UMP level α, then ϕ is unbiased level α.*
(iii) *If β_ϕ continuous for each ϕ and ϕ_0 is UMP level α and α-similar then ϕ_0 is UMPU.*

Proof. (i) $\beta_\phi \leq \alpha$ on Ω_H, $\beta_\phi \geq \alpha$ on Ω_A and β_ϕ continuous implies $\beta_\phi = \alpha$ on $\overline{\Omega}_H \cap \overline{\Omega}_A$.
(ii) Let $\phi^\alpha \equiv \alpha$. Since ϕ is UMP $\beta_\phi \geq \beta_{\psi^\alpha} = \alpha$ on Ω_A. Hence ϕ is unbiased level α.
(iii) Since ϕ^α is α-similar and ϕ_0 is UMP among α-similar tests we have $\beta_{\phi_0} \geq \beta_{\psi^\alpha} = \alpha$ on Ω_A. Hence ϕ_0 is unbiased level α. By continuity of β_ϕ any α-similar level α test ϕ is unbiased level α so $\beta_{\phi_0} \geq \beta_\phi$ on Ω_A. Thus ϕ_0 is UMPU. □

Theorem 27. *Consider a one parameter exponential family with its natural parameter and the hypothesis $H : \Theta \in [\theta_1, \theta_2]$ vs $A : \Theta \notin [\theta_1, \theta_2]$, $\theta_1 < \theta_2$. Let ϕ be any test of H vs A. Then there is a test ψ of the form*

$$\psi(x) = \begin{cases} 1, & x \notin (c_1, c_2), \\ \gamma_i, & x = c_i, \\ 0, & x \in (c_1, c_2), \end{cases}$$

such that $\beta_\psi(\theta_i) = \beta_\phi(\theta_i)$, $\beta_\psi(\theta) \leq \beta_\phi(\theta)$ on Ω_H and $\beta_\psi(\theta) \geq \beta_\phi(\theta)$ on Ω_A. Moreover, if $\beta_\psi(\theta_i) = \alpha$, then ψ is UMPU level α.

Proof. Put $\alpha_i = \beta_\phi(\theta_i)$. One can find a test ϕ_0 of the form in Theorem 3, Lecture 15, such that $\beta_{\phi_0}(\theta_i) = 1 - \alpha_i$ (we have not proved this in class, see Lemma 4.81, p. 248) and then this ϕ_0 minimizes the power function on $(\infty, \theta_1) \cup (\theta_2, \infty)$ and maximizes it on (θ_1, θ_2) among all tests ϕ' subject to $\beta_{\phi'}(\theta_i) = 1 - \alpha_i$. But then, $\psi = 1 - \phi_0$ satisfies $\beta_\psi(\theta_i) = \alpha_i$ and maximizes the power function on $(\infty, \theta_1) \cup (\theta_2, \infty)$ and minimizes it on (θ_1, θ_2) among all test subject to the restrictions. This proves the first part.

If $\beta_\psi(\theta_i) = \alpha$, then ψ is α-similar and hence ψ is UMP level α among all α-similar tests. For a one parameter exponential family β_ϕ is continuous for all ϕ so (iii) in the Proposition shows that ψ is UMPU level α. □

Point hypothesis. In this section we are concerned with hypothesis of the form $H : \Theta = \theta_0$ vs $A : \Theta \neq \theta_0$. Again it seems reasonable that tests of the form ψ in Theorem 27 are appropriate.

Theorem 28. *Consider a one parameter exponential family with natural parameter and $\Omega_H = \{\theta_0\}$, $\Omega_A = \Omega \setminus \{\theta_0\}$ where θ_0 is in the interior of Ω. Let ϕ be any test of H vs A. Then there is a test of the form ψ in Theorem 27 such that*

$$\begin{aligned} \beta_\psi(\theta_0) &= \beta_\phi(\theta_0), \\ \partial_\theta \beta_\psi(\theta_0) &= \partial_\theta \beta_\phi(\theta_0) \end{aligned} \tag{18.2}$$

and for $\theta \neq \theta_0$, $\beta_\psi(\theta)$ is maximized among all tests satisfying the two equalities. Moreover, If ψ has size α and $\partial \beta_\psi(\theta_0) = 0$, then it is UMPU level α.

Sketch of proof. First one need to show that there are tests of the form ψ that satisfies the equialities.

Put $\alpha = \beta_\phi(\theta_0)$ and $\gamma = \partial_\theta \beta_\phi(\theta_0)$. Let ϕ_u be the UMP level u test for testing $H : \Theta \geq \theta_0$ vs $A : \Theta < \theta_0$, and for $0 \leq u \leq \alpha$ put

$$\phi'_u(x) = \phi_u(x) + 1 - \phi_{1-\alpha+u}(x).$$

Note that, for each $0 \leq u \leq \alpha$,

$$\beta_{\phi'_u}(\theta_0) = \beta_{\phi_u}(\theta_0) + 1 - \beta_{\phi_{1-\alpha+u}}(\theta_0) = u + 1 - (1 - \alpha + u) = \alpha.$$

Then ϕ'_u has the right form, i.e. as in Theorem 27. The test $\phi'_0 = 1 - \phi_{1-\alpha}$ has level α and is by construction UMP level α for testing $H' : \Theta = \theta_0$ vs $A' : \Theta > \theta_0$. Similarly $\phi'_\alpha = \phi_\alpha$ is UMP level α for testing $H' : \Theta = \theta_0$ vs $A'' : \Theta < \theta_0$. We claim that

(i) $\partial_\theta \beta_{\phi'_\alpha}(\theta_0) \le \gamma \le \partial_\theta \beta_{\phi'_0}(\theta_0)$.
(ii) $u \mapsto \partial_\theta \beta_{\phi_u}(\theta_0)$ is continuous.

The first is easy to see intuitively in a picture. A complete argument is in Lemma 4.103, p. 257 in Schervish. The second is a bit involved and we omitt it here. See p. 259 for details. From (i) and (ii) we conclude that there is a test of the form ψ (i.e. ϕ'_{u_0} for some u_0) that satisfies (18.2).

It remains to show that this test maximizes the power function among all level α tests satisfying the restriction on the derivative. For any test η we have

$$\begin{aligned}\partial_\theta \beta_\eta(\theta_0) &= \partial_\theta \int_{\mathcal{X}} \eta(x)c(\theta)e^{\theta x}\nu(dx)|_{\theta=\theta_0} \\ &= \int_{\mathcal{X}} \eta(x)(c(\theta_0)x + c'(\theta_0))e^{\theta_0 x}\nu(dx) \\ &= E_{\theta_0}[X\eta(X)] - \beta_\eta(\theta_0)E_{\theta_0}[X],\end{aligned}$$

where we used integration by parts in the last step. Hence, $\partial_\theta \beta_\eta(\theta_0) = \gamma$ iff

$$E_{\theta_0}[X\eta(X)] = \gamma + \alpha E_{\theta_0}[X].$$

Note that the RHS does not depend on η. For any $\theta_1 \neq \theta_0$ and put

$$\begin{aligned}p_0(x) &= c(\theta_1)e^{\theta_1 x} \\ p_1(x) &= c(\theta_0)e^{\theta_0 x} \\ p_2(x) &= xc(\theta_0)e^{\theta_0 x}.\end{aligned}$$

Then

$$E_{\theta_0}[X\eta(X)] = \int \eta(x)p_2(x)\nu(dx)$$

We know from last time (or Lemma 4.78, p. 247 using Lagrange multipliers) that a test of the form

$$\eta_0(x) = \begin{cases} 1, & p_0(x) > \sum_{i=1}^2 k_i p_i(x), \\ \gamma(x), & p_0(x) = \sum_{i=1}^2 k_i p_i(x), \\ 0, & p_0(x) < \sum_{i=1}^2 k_i p_i(x), \end{cases}$$

where $0 \le \gamma(x) \le 1$ and k_i are constants, maximizes $\int \eta(x)p_0(x)\nu(dx)$ subject to the constraints

$$\begin{aligned}\int \eta(x)p_i(x)\nu(dx) &\le \int \eta_0(x)p_i(x)\nu(dx), \text{ for } i \text{ such that } k_i > 0, \\ \int \eta(x)p_i(x)\nu(dx) &\ge \int \eta_0(x)p_i(x)\nu(dx), \text{ for } i \text{ such that } k_i < 0.\end{aligned}$$

That is, it maximizes $\beta_\eta(\theta_1)$ subject to

$$\begin{aligned}\beta_\eta(\theta_0) &\le (\ge) \beta_{\eta_0}(\theta_0) \\ \mathrm{E}_{\theta_0}[\eta(X)] &\le (\ge)\, \mathrm{E}_{\theta_0}[\eta_0(X)],\end{aligned}$$

where the direction of the inequalities depend on k_i.

The test η_0 corresponds to rejecting the hypothesis if

$$e^{(\theta_1-\theta_0)x} > k_1 + k_2 x.$$

By choosing k_1 and k_2 appropriety we can get a test of the form ψ which is the same for all $\theta_1 \neq \theta_0$.

Finally, we want to show that if the test is level α and $\partial_\theta \beta_\phi(\theta_0) = 0$, the the test is UMPU level α. For this we only need to show that $\partial_\theta \beta_\phi(\theta_0) = 0$ is necessary for the test to be unbiased. But this is obvious because, since the power function is differentiable, if the derivative is either strictly positive or strictly then the power function is less than α in some left- or right-neighborhood of θ_0. □

19. Nuisance parameters

Suppose the parameter Θ is multidimensional $\Theta = (\Theta_1, \dots, \Theta_k)$ and Ω_H is of lower dimension than k, say d dimensional $d < k$, then the remaining parameters are called *nuisance parameters*.

Let $\mathcal{P}_0$ be a parametric family $\mathcal{P}_0 = \{P_\theta : \theta \in \Omega\}$. Let $G \subset \Omega$ be a subparameter space and $\mathcal{Q}_0 = \{P_\theta : \theta \in G\}$ be a subfamily of $\mathcal{P}_0$. Let Ψ be the parameter of the family $\mathcal{Q}_0$.

Definition 33. If T is a sufficient statistic for Ψ in the classical sense, then a test ϕ has *Neyman structure relative to G and T* if $E_\theta[\phi(X) \mid T = t]$ is constant as a function of t P_θ-a.s. for all $\theta \in G$.

Why is Neyman structure a good thing? It is because it sometimes enables a procedure to obtain UMPU tests. Suppose that we can find statistic T such that the distribution of X conditional on T has one-dimensional parameter. Then we can try to find the UMPU test among all tests that have level α conditional on T. Then this test will also be UMPU level α unconditionally.

There is a connection here with α-similar tests.

Lemma 6. *If H is a hypothesis and $\mathcal{Q}_0 = \{P_\theta : \theta \in \overline{\Omega}_H \cap \overline{\Omega}_A\}$ and ϕ has Neyman structure, then ϕ is α-similar.*

Proof. Since

$$\beta_\phi(\theta) = E_\theta[\phi(X)] = E_\theta[E_\theta[\phi(X) \mid T]]$$

and $E_\theta[\phi(X) \mid T]$ is constant for $\theta \in \overline{\Omega}_H \cap \overline{\Omega}_A$ we see that $\beta_\phi(\theta)$ is constant on $\overline{\Omega}_H \cap \overline{\Omega}_A$. □

There is a converse under some slightly stronger assumptions.

Lemma 7. *If T is a boundedly complete sufficient statistic for the subparameter space $G \subset \Omega$, then every α-similar test on G has Neyman structure relative to G and T.*

Proof. By α-similarity $E_\theta[E[\phi(X) \mid T] - \alpha] = 0$ for all $\theta \in G$. Since T is boundedly complete we must have $E[\phi(X) \mid T] = \alpha$ P_θ-a.s. for all $\theta \in G$. □

Now we can use this to find conditions when UMPU tests exists.

Proposition 5. *Let $G = \overline{\Omega}_H \cap \overline{\Omega}_A$. Let I be an index set such that $G = \cup_{i \in I} G_i$ is a partition of G. Suppose there exists a statistic T that is boundedly complete sufficient statistic for each subparameter space G_i. Assume that the power function* *of every test is continuous. If there is a UMPU level α test ϕ among those which have Neyman structure relative to G_i and T for all $i \in I$, then ϕ is UMPU level α.*

Proof. From last time (Proposition 4(i)) we know that continuity of the power function implies that all unbiased level α tests are α-similar. By the previous lemma every α-similar test has Neyman structure. Since ϕ is UMPU level α among all such tests it is UMPU level α. □

In the case of exponential families one can prove the following.

Theorem 29. *Let $X = (X_1, \dots, X_k)$ have a k-parameter exponential family with $\Theta = (\Theta_1, \dots, \Theta_k)$ and let $U = (X_2, \dots, X_k)$.*

(i) *Suppose that the hypothesis is one-sided or two-sided concerning only Θ_1. Then there is a UMP level α test conditional on U, and it is UMPU level α.*
(ii) *If the hypothesis concerns only Θ_1 and the alternative is two-sided, then there is a UMPU level α test conditional on U, and it is also UMPU level α.*

Proof. Suppose that the density is

$$f_{X|\Theta}(x \mid \theta) = c(\theta)h(x)\exp\{\sum_{i=1}^{k} \theta_i x_i\}.$$

Let $G = \overline{\Omega}_H \cap \overline{\Omega}_A$. The conditional density of X_1 given $U = u = (x_1, \dots, x_k)$ is

$$f_{X_1|\Theta,U}(x_1 \mid \theta, u) = \frac{c(\theta)h(x)e^{\sum_{i=1}^{k} \theta_i x_i}}{\int c(\theta)h(x)e^{\sum_{i=1}^{k} \theta_i x_i} dx_1} = \frac{h(x)e^{\theta_1 x_1}}{\int h(x)e^{\theta_1 x_1} dx_1}.$$

This is a one-parameter exponential family with natural parameter Θ_1.

For the hypothesis (one- or two-sided) we have that G is either $G_0 = \{\theta : \theta_1 = \theta_1^0\}$ some θ_1^0 or the union $G_1 \cup G_1$ with $G_1 = \{\theta : \theta_1 = \theta_1^1\}$, $G_2 = \{\theta : \theta_1 = \theta_1^2\}$. The parameter $\Psi = (\Theta_2, \dots, \Theta_k)$ has a complete sufficient statistic $U = (X_2, \dots, X_k)$.

Let η be an unbiased level α test. Then by Proposition 4(i), η is α-similar on G_0, G_1, and G_2. By the previous lemma η has Neyman structure. Moreover, for every test η, $\beta_\eta(\theta) = E_\theta[E_\theta[\eta(X) \mid U]]$ so a test that maximizes the conditional power function uniformly for $\theta \in \Omega_A$ subject to contraints also maximizes the marginal power function subject to the same contstraints.

For part (i) in the conditional problem given $U = u$ there is a level α test that maximizes the conditional power function uniformly on Ω_A subject to having Neyman structure. Since every unbiased level α test has Neyman structure and the power function is the expectation of the conditional power function ϕ is UMPU level α.

For part (ii), if $\Omega_H = \{\theta : c_1 \le \theta_1 \le c_2\}$ with $c_1 < c_2$, then as above the conditional UMPU level α test ϕ is also UMPU level α.

For a point hypothesis $\Omega_H = \{\theta : \theta_1 = \theta_1^0\}$ we must take partial derivative of $\beta_\eta(\theta)$ with respect to θ_1 at every point in G. A little more work... □