ON THE POWER FUNCTION OF THE ANALYSIS OF VARIANCE TEST[1]

By Abraham Wald

Columbia University

It is known[2] that the general problem of the analysis of variance can be reduced by an orthogonal transformation to the following canonical form: Let the variates $y_1, \cdots, y_p, z_1, \cdots, z_n$ be independently and normally distributed with a common unknown variance σ^2. The mean values of $z_1, \cdots, z_n$ are known to be zero, and the mean values $\eta_1, \cdots, \eta_p$ of the variates $y_1, \cdots, y_p$ are unknown. The canonical form of the analysis of variance test is the test of the hypothesis that

$$\eta_1 = \eta_2 = \cdots = \eta_r = 0 \qquad (r \leq p) \tag{1}$$

where a single observation is made on each of the variates $y_1, \cdots, y_p, z_1, \cdots, z_n$.

In the theory of the analysis of variance the test of the hypothesis (1) is based on the critical region

$$\frac{y_1^2 + \cdots + y_r^2}{z_1^2 + \cdots + z_n^2} \geq c \tag{2}$$

where the constant c is chosen so that the size of the critical region is equal to the level of significance α we wish to have. The critical region (2) is identical with the critical region

$$\frac{y_1^2 + \cdots + y_r^2}{y_1^2 + \cdots + y_r^2 + z_1^2 + \cdots + z_n^2} \geq c' = \frac{c}{c+1}. \tag{3}$$

It is known that the power function of the critical region (3) depends only on the single parameter

$$\lambda = \frac{1}{\sigma^2} \sum_{i=1}^{r} \eta_i^2. \tag{4}$$

Denote the power function of the critical region (3) by $\beta_0(\lambda)$. P. L. Hsu has proved[3] the following optimum property of the region (3): *Let W be a critical region which satisfies the following two conditions*:

(a) *The size of W is equal to the size of the region* (3).

---

[1] Presented at a joint meeting of the Institute of Mathematical Statistics and the American Mathematical Society in New York, December, 1941.

[2] See for instance P. C. Tang, "The power function of the analysis of variance tests," *Stat. Res. Mem.*, Vol. 2, 1938.

[3] P. L. Hsu, "Analysis of variance from the power function standpoint," *Biometrika*, January, 1941.

(b) *The power function of W depends on the single parameter λ.*
Then $\beta(\lambda) \leq \beta_0(\lambda)$ where $\beta(\lambda)$ denotes the power function of W.

Condition (b) is a serious restriction in Hsu's result. In this paper we shall prove an optimum property of $\beta_0(\lambda)$ where $\beta_0(\lambda)$ is compared with the power function of any other critical region of size equal to that of (3).

For any given values $\eta'_{r+1}, \cdots, \eta'_p, \sigma'$ and λ denote by $S(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda)$ the sphere defined by the equations

$$(5) \qquad \eta_1^2 + \cdots + \eta_r^2 = \lambda\sigma'^2; \qquad \eta_i = \eta'_i (i = r+1, \cdots, p); \qquad \sigma = \sigma'.$$

For any region W denote by $\beta_W(\eta_1, \cdots, \eta_p, \sigma)$ the power function of W, i.e. $\beta_W(\eta_1, \cdots, \eta_p, \sigma)$ denotes the probability that the sample point will fall within W calculated under the assumption that $\eta_1, \cdots, \eta_p$ and σ are the true values of the parameters. We will denote by $\gamma_W(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda)$ the integral of the power function $\beta_W(\eta'_1, \cdots, \eta'_p, \sigma')$ over the surface $S(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda)$ divided by the area of $S(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda)$, i.e.

$$(6) \qquad \begin{aligned} &\gamma_W(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda) \\ &\quad = \left[\int_{S(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda)} dA\right]^{-1} \int_{S(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda)} \beta_W(\eta'_1, \cdots, \eta'_p, \sigma')\, dA. \end{aligned}$$

We will prove the following

THEOREM: *If W is a critical region of size equal to that of* (3), *i.e.* $\beta_W(0, \cdots, 0, \eta_{r+1}, \cdots, \eta_p, \sigma) = \beta_0(0)$, *then*

$$(7) \qquad \gamma_W(\eta'_{r+1}, \cdots, \eta'_p, \sigma', \lambda) \leq \beta_0(\lambda)$$

for arbitrary values $\eta'_{r+1}, \cdots, \eta'_p, \sigma'$ and λ.

If W satisfies Hsu's condition (*b*) then the power function $\beta_W(\eta_1, \cdots, \eta_p, \sigma)$ is constant on the surface $S(\eta_{r+1}, \cdots, \eta_p, \sigma, \lambda)$ and therefore $\gamma_W(\eta_{r+1}, \cdots, \eta_p, \sigma, \lambda) = \beta_W(\eta_1, \cdots, \eta_p, \sigma)$. Hence Hsu's result is an immediate consequence of our Theorem.

Denote $|\sqrt{y_1^2 + \cdots + y_r^2 + z_1^2 + \cdots + z_n^2}|$ by t and for any values $a_{r+1}, \cdots, a_p, b$ let $R(a_{r+1}, \cdots, a_p, b)$ be the set of all sample points for which

$$y_i = a_i (i = r+1, \cdots, p) \quad \text{and} \quad t = b.$$

For any region W of the sample space we denote by $W(y_{r+1}, \cdots, y_p, t)$ the common part of W and $R(y_{r+1}, \cdots, y_p, t)$.

In order to prove our Theorem we first show the validity of the following

LEMMA 1: *For any critical region Z there exists a function $\varphi_Z(y_{r+1}, \cdots, y_p, t)$ of the variables $y_{r+1}, \cdots, y_p, t$ such that the critical region Z^* defined by the inequality*

$$y_1^2 + \cdots + y_r^2 \geq \varphi_Z(y_{r+1}, \cdots, y_p, t)$$

satisfies the following two conditions:

$$\text{(a)} \qquad \beta_Z(0, \cdots, 0, \eta_{r+1}, \cdots, \eta_p, \sigma) = \beta_{Z^*}(0, \cdots, 0, \eta_{r+1}, \cdots, \eta_p, \sigma);$$

(b) $$\gamma_Z(\eta_{r+1}, \cdots, \eta_p, \sigma, \lambda) \leq \gamma_{Z^*}(\eta_{r+1}, \cdots, \eta_p, \sigma, \lambda).$$

PROOF: Denote by $P_Z(y_{r+1}, \cdots, y_p, t)$ the conditional probability of $Z(y_{r+1}, \cdots, y_p, t)$ calculated under the condition that the sample point lies in $R(y_{r+1}, \cdots, y_p, t)$ and under the assumption that $\eta_1 = \cdots = \eta_r = 0$. Denote by $F(d, t)$ the conditional probability that

$$y_1^2 + \cdots + y_r^2 \geq d$$

calculated under the condition that the sample point lies in $R(y_{r+1}, \cdots, y_p, t)$ and under the assumption that $\eta_1 = \cdots = \eta_r = 0$. It is easy to verify that the values of $F(d, t)$ and $P_Z(y_{r+1}, \cdots, y_p, t)$ do not depend on the unknown parameters $\eta_{r+1}, \cdots, \eta_p, \sigma$. Since $F(d, t)$ is a continuous function of d and since $F(t^2, t) = 0$, there exists a function $\varphi_Z(y_{r+1}, \cdots, y_p, t)$ such that

$$F[\varphi_Z(y_{r+1}, \cdots, y_p, t), t] = P_Z(y_{r+1}, \cdots, y_p, t).$$

For this function $\varphi_Z(y_{r+1}, \cdots, y_p, t)$ the region Z^* certainly satisfies condition (a) of Lemma 1. We will show that condition (b) is also satisfied. Consider the ratio

$$\frac{\int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{p}(y_i - \eta_i)^2 - \frac{1}{2\sigma^2}\sum_{\alpha=1}^{n} z_\alpha^2\right] dA}{\exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{r} y_i^2 + \sum_{i=r+1}^{p}(y_i - \eta_i)^2 + \sum_{\alpha=1}^{n} z_\alpha^2\right)\right]} \tag{8}$$

$$= e^{-\frac{1}{2}\lambda} \int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} e^{\sum_{i=1}^{r} y_i\eta_i/\sigma^2} \, dA.$$

Denote $\left|\sqrt{\sum_{i=1}^{r} y_i^2}\right|$ by r_y. Then we have

$$\int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} e^{\sum_{i=1}^{r} y_i\eta_i/\sigma^2} \, dA = \int_{(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} e^{\sqrt{\lambda}\, r_y \cos[\alpha(\eta)]/\sigma} \, dA, \tag{9}$$

where $\alpha(\eta)$ denotes the angle $(0 \leq \alpha(\eta) \leq \pi)$ between the vector y with the components $y_1, \cdots, y_r$ and the vector η with the components $\eta_1, \cdots, \eta_r$. Because of the symmetry of the sphere, the value of the right hand side of (9) is not changed if we substitute $\beta(\eta)$ for $\alpha(\eta)$ where $\beta(\eta)$ denotes the angle $(0 \leq \beta(\eta) \leq \pi)$ between the vector η and an arbitrarily chosen fixed vector u. Hence the value of the right hand side of (9) depends only on r_y, i.e.

$$\int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} e^{\sqrt{\lambda}\, r_y \cos[\alpha(\eta)]/\sigma} \, dA \tag{10}$$

$$= \int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} e^{\sqrt{\lambda}\, r_y \cos[\beta(\eta)]/\sigma} \, dA = I(r_y).$$

Now we will show that $I(r_y)$ is a monotonically increasing function of r_y. We have

$$(11)\qquad \frac{dI(r_y)}{dr_y} = \frac{\sqrt{\lambda}}{\sigma}\int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} \cos[\beta(\eta)]\, e^{\sqrt{\lambda}\, r_y \cos[\beta(\eta)]/\sigma}\, dA.$$

Denote by ω_1 the subset of $S(\eta_{r+1}, \cdots, \eta_p, \sigma, \lambda)$ in which $0 \leq \beta(\eta) \leq \frac{\pi}{2}$ and by ω_2 the subset in which $\frac{\pi}{2} \leq \beta(\eta) \leq \pi$. Because of the symmetry of the sphere we obviously have

$$(12)\qquad \begin{aligned} \int_{\omega_2} \cos[\beta(\eta)] e^{\sqrt{\lambda}\, r_y \cos[\beta(\eta)]/\sigma}\, dA &= \int_{\omega_1} \cos[\pi - \beta(\eta)] e^{\sqrt{\lambda}\, r_y \cos[\pi-\beta(\eta)]/\sigma}\, dA \\ &= -\int_{\omega_1} \cos[\beta(\eta)] e^{-\sqrt{\lambda}\, r_y \cos[\beta(\eta)]/\sigma}\, dA. \end{aligned}$$

Hence

$$(13)\qquad \frac{dI(r_y)}{dr_y} = \frac{\sqrt{\lambda}}{\sigma}\int_{\omega_1} \cos[\beta(\eta)]\{e^{\sqrt{\lambda}\, r_y \cos[\beta(\eta)]/\sigma} - e^{-\sqrt{\lambda}\, r_y \cos[\beta(\eta)]/\sigma}\}\, dA.$$

The right hand side of (13) is positive. Hence $I(r_y)$, and therefore also the left hand side of (8), is a monotonically increasing function of r_y.

Let $P_1(y'_{r+1}, \cdots, y'_p, t', \eta_1, \cdots, \eta_p, \sigma)\, dy_{r+1} \cdots dy_p\, dt$ be the probability that the sample point will fall in the intersection of Z and the set

$$y'_i - \tfrac{1}{2}\, dy_i \leq y_i \leq y'_i + \tfrac{1}{2}\, dy_i (i = r + 1, \cdots, p), \quad t' - \tfrac{1}{2}\, dt \leq t \leq t' + \tfrac{1}{2}\, dt$$

Similarly let $P_2(y'_{r+1}, \cdots, y'_p, t', \eta_1, \cdots, \eta_p, \sigma)\, dy_{r+1} \cdots dy_p\, dt$ be the unconditional probability that the sample point will fall in the intersection of Z^* and the set

$$y'_i - \tfrac{1}{2}\, dy_i \leq y_i \leq y'_i + \tfrac{1}{2}\, dy_i (i = r + 1, \cdots, p), \quad t' - \tfrac{1}{2}\, dt \leq t \leq t' + \tfrac{1}{2}\, dt.$$

Since the function $\varphi_Z(y_{r+1}, \cdots, y_p, t)$ has been defined so that

$$P_Z(y_{r+1}, \cdots, y_p, t) = F[\varphi(y_{r+1}, \cdots, y_p, t), t],$$

we obviously have

$$(14)\qquad \begin{aligned} &P_1(y_{r+1}, \cdots, y_p, t, 0, \cdots, 0, \eta_{r+1}, \cdots, \eta_p, \sigma) \\ &\qquad = P_2(y_{r+1}, \cdots, y_p, t, 0, \cdots, 0, \eta_{r+1}, \cdots, \eta_p, \sigma). \end{aligned}$$

Using a lemma[4] by Neyman and Pearson, we easily obtain

$$(15)\qquad \begin{aligned} &\int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} P_2(y_{r+1}, \cdots, y_p, t, \eta_1, \cdots, \eta_p, \sigma)\, dA \\ &\qquad \geq \int_{S(\eta_{r+1},\cdots,\eta_p,\sigma,\lambda)} P_1(y_{r+1}, \cdots, y_p, t, \eta_1, \cdots, \eta_p, \sigma)\, dA \end{aligned}$$

---

[4] J. Neyman and E. S. Pearson, "Contributions to the theory of testing statistical hypotheses," *Stat. Res. Mem.*, Vol. 1, London, 1936.

from (14) and the fact that the left hand side of (8) is a monotonically increasing function of $r_y^2 = y_1^2 + \cdots + y_r^2$. Condition (b) is an immediate consequence of (15). Hence Lemma 1 is proved.

For the proof of our theorem we will also need the following

LEMMA 2: *Let $v_1, \cdots, v_k$ be k normally and independently distributed variates with a common variance σ^2. Denote the mean value of v_i by $\alpha_i (i = 1, \cdots, k)$ and let $f(v_1, \cdots, v_k, \sigma)$ be a function of the variables $v_1, \cdots, v_k$ and σ which does not involve the mean values $\alpha_1, \cdots, \alpha_k$. Then, if the expected value of $f(v_1, \cdots, v_k, \sigma)$ is equal to zero, $f(v_1, \cdots, v_k, \sigma)$ is identically equal to zero, except perhaps on a set of measure zero.*

PROOF: Lemma 2 is obviously proved for all values of σ if we prove it for $\sigma = 1$. Hence we will assume that $\sigma = 1$. It is known that a k-variate distribution which has moments equal to those of the joint distribution of $v_1, \cdots, v_k$, must be identical with the joint distribution of $v_1, \cdots, v_k$. That is to say, the joint distribution of $v_1, \cdots, v_k$ is uniquely determined by its moments. Hence if

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} v_1^{r_1} v_2^{r_2} \cdots v_k^{r_k} g(v_1, \cdots, v_k) e^{-\frac{1}{2}\sum_{i=1}^{k}(v_i-\alpha_i)^2} dv_1 \cdots dv_k = 0 \tag{16}$$

for any set $(r_1, \cdots, r_k)$ of non-negative integers, then $g(v_1, \cdots, v_k)$ must be equal to zero except perhaps on a set of measure zero. Now let $f(v_1, \cdots, v_k)$ be a function whose expected value is zero, i.e.

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(v_1, \cdots, v_k) e^{-\frac{1}{2}\sum_{i=1}^{k}(v_i-\alpha_i)^2} dv_1 \cdots dv_k = 0 \tag{17}$$

identically in $\alpha_1, \cdots, \alpha_k$. From (17) it follows that

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(v_1, \cdots, v_k) e^{-\frac{1}{2}\sum_{i=1}^{k} v_i^2 + \sum_{i=1}^{k} \alpha_i v_i} dv_1 \cdots dv_k = 0 \tag{18}$$

identically in $\alpha_1, \cdots, \alpha_k$. Differentiating the left hand side of (18) r_1 times with respect to α_1, r_2 times with respect to α_2, $\cdots$, and r_k times with respect to α_k, we obtain

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} v_1^{r_1} \cdots v_k^{r_k} f(v_1, \cdots, v_k) e^{-\frac{1}{2}\sum(v_i-\alpha_i)^2} dv_1 \cdots dv_k = 0. \tag{19}$$

From (16) and (19) it follows that $f(v_1, \cdots, v_k) = 0$. Hence Lemma 2 is proved.

Using Lemmas 1 and 2 we can easily prove our theorem. Because of Lemma 1 we can restrict ourselves to critical regions W which are given by an inequality of the following type

$$y_1^2 + \cdots + y_r^2 \geq \varphi(y_{r+1}, \cdots, y_p, t)$$

where $\varphi(y_{r+1}, \cdots, y_p, t)$ is some function of $y_{r+1}, \cdots, y_p$ and t. The above inequality can be written as

$$\frac{y_1^2 + \cdots + y_r^2}{t^2} \geq \psi(y_{r+1}, \cdots, y_p, t). \tag{20}$$

For any given values of $y_{r+1}, \cdots, y_p, t$ denote by $P(y_{r+1}, \cdots; y_p, t)$ the conditional probability that (20) holds calculated under the assumption that $\eta_1 = \cdots = \eta_r = 0$. It is obvious that $P(y_{r+1}, \cdots, y_p, t)$ does not depend on the unknown parameters $\eta_{r+1}, \cdots, \eta_p, \sigma$. If we denote by W the critical region defined by the inequality (20), we have

$$\begin{aligned} &\beta_W(0, \cdots, 0, \eta_{r+1}, \cdots, \eta_p, \sigma) \\ (21) \quad &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_0^{\infty} P(y_{r+1}, \cdots, y_p, t)\rho_1(y_{r+1}, \cdots, y_p, \eta_{r+1}, \cdots, \eta_p, \sigma) \\ &\qquad \times \rho_2(t, \sigma)\, dy_{r+1} \cdots dy_p\, dt \end{aligned}$$

where $\rho_1(y_{r+1}, \cdots, y_p, \eta_{r+1}, \cdots, \eta_p, \sigma)$ denotes the joint probability density function of $y_{r+1}, \cdots, y_p$ and $\rho_2(t, \sigma)$ denotes the probability density function of t calculated under the assumption that $\eta_1 = \cdots = \eta_r = 0$. In order to satisfy the condition of our Theorem, the function ψ in (20) must be chosen so that

$$\begin{aligned} (22) \quad &\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_0^{\infty} P(y_{r+1}, \cdots, y_p, t)\rho_1(y_{r+1}, \cdots, y_p, \eta_{r+1}, \cdots, \eta_p, \sigma) \\ &\qquad \times \rho_2(t, \sigma)\, dy_{r+1} \cdots dy_p\, dt = \beta_0(0). \end{aligned}$$

Let

$$\int_0^{\infty} P(y_{r+1}, \cdots, y_p, t)\rho_2(t, \sigma)\, dt = Q(y_{r+1}, \cdots, y_p, \sigma). \tag{23}$$

Then we obtain from (22)

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} Q(y_{r+1}, \cdots, y_p, \sigma)\rho_1\, dy_{r+1} \cdots dy_p = \beta_0(0). \tag{24}$$

From (24) and Lemma 2 it follows that

$$Q(y_{r+1}, \cdots, y_p, \sigma) = \beta_0(0) \tag{25}$$

except perhaps on a set of measure zero. From (23), (25) and a result[5] by P. L. Hsu we obtain

$$P(y_{r+1}, \cdots, y_p, t) = \beta_0(0) \tag{26}$$

except perhaps on a set of measure zero.

It follows easily from (26) that $\psi(y_{r+1}, \cdots, y_p, t)$ is equal to a fixed constant except perhaps on a set of measure zero. This proves our Theorem.

[5] P. L. Hsu, "Notes on Hotelling's generalized T," *Annals of Math. Stat.*, Vol. 9, p. 237.