BULLETIN OF THE
GREEK MATHEMATICAL SOCIETY
Volume 57, 2010 (251–264)

Adopting Hypergeometric Functions for Sequential Statistical Methods

C. P. Kitsos

Abstract

The Hypergeometric Functions are, in principle, applied to Mathematical Physics, solving differential equations. In this paper we discuss their application to Statistics and we introduce applications to sequential statistics due to Sequential Probability Ratio Test (SPRT). The sequential principle plays an important role to optimal experiment design theory. Therefore we provide evidence that the Hypergeometric Functions play an important hidden role to Statistics.

Keywords: Hypergeometric functions, Sequential Probability Ratio Test (SPRT), Least Square Method.

1. Introduction

In his 1812 pioneering paper (Disquisitiones Generales Circa seriem infinitam $1 + \frac{\alpha \cdot \beta}{\gamma \cdot 1} + \ldots$) Gauss introduced that series, which latter J.K. Plaff called hypergeometric series.

There is a number of applications of the hypergeometric functions (HF) which as briefly refered here but the emphasis will be for the multivariate sequential statistical problems. We discuss the statistical implimentations, considering the HF as the main mathematical tool to support the Sequential Probability Ratio Test, (SPRT), Kitsos (2007) [12], for the statistical problems. Firstly let us consider the function

$$h_{1/2}(\alpha; z) = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{\alpha+r}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \cdot \frac{z^r}{r!} \ \ with \ \ z \in \mathbb{R}, \alpha > 0 \ .$$

The function $h_{1/2}(\alpha; z)$ defines the non-central t distribution with n degrees of freedom and non-centrality parameter $\delta \in \mathbb{R}$, Graybill (1976) [8] among others. Secondly if we define the function

$$h_1(\alpha; z) = \sum_{r=0}^{\infty} \frac{\Gamma(\alpha)}{\Gamma(\alpha + r)} \cdot \frac{z^r}{r!} \ \ with \ \ z \geq 0, \alpha > 0,$$

then, the distribution function of the non-central $\mathcal{X}_n^2(\delta), \delta > 0$ can be considered. The confluent hypergeometric function "extending" $h_1(\alpha; z)$ is defined as

$$H_{1,1}(\alpha;\beta;z) = \sum_{r=0}^{\infty} \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} \cdot \frac{\Gamma(\beta)}{\Gamma(\beta+r)} \cdot \frac{z^r}{r!} \ \ with\ \ z \geq 0, \ \ \alpha, \beta > 0,$$

and we can define the non-central distribution of $F_{m,n}(\delta), \delta \geq 0$ due to it, see Examples 1.2, 1.3 below. Eventually the hypergeometric function (HF) is defined as

$$H_{2,1}(\alpha,\beta;\gamma;z) = \sum_{r=0}^{\infty} \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} \cdot \frac{\Gamma(\beta+r)}{\Gamma(\beta)} \cdot \frac{\Gamma(\gamma)}{\Gamma(\gamma+r)} \cdot \frac{z^r}{r!} \ \ with\ \ |z| < 1,$$

and can be applied to various sequential stoctistical tests: $t, F, \mathcal{X}^2, T^2$ (see also Kitsos (1993) [11]) when invariant Sequential Probability Ratio Test (SPRT) are considered. That is this paper bridges the HF with SPRT of the statistical line of thought, and we shall refer to these tests in sections 3 and 4.

It is known that the Lagrage, Chebyshev and Legendue polynomials, are applied to statistics, throught the optimal design theory, Pukelsheim (1993) [19], and these polynomials are related to hypergeometric series. Therefore we link relations in Proposition 1.1 extending the existend statistics ideas. The question we try to answer in this paper is: how the hypergeometric functions are applied to statistics and especially to Sequential Analysis.

The Laguerre polynomials, can be also defined in terms of the confluent hypergeometric functions. Moreover Fisher's information matrix can be evaluated for the class of orthogonal polynomials (defined through the second order differential operator) of hypergeometric type, see Sanchez-Ruiz and Dehesa (2005) [21].

Example 1.1 Let $X_i, \ i = 1, 2, \ldots, n$ be independent random variables from the normal distribution ie $X_i \sim N\left(\mu_i, \sigma_i^2\right)$. Then $Z_i = \frac{x_i - \mu_i}{\sigma_i} \sim N(0,1)$. Let us define

$$X_*^2 = \sum_{i=1}^{n} \left(\frac{x_i}{\sigma_i}\right)^2 \ \ and\ \ t_* = \frac{\overline{x}_n}{S_n/\sqrt{n}-1}\ , \ \ with$$

$$\overline{x}_n = n^{-1} \sum_{i=1}^{n} x_i,\ S_n^2 = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 \quad .$$

Then X_*^2 follows the non-central chi-square distribution with non-centrality parameter φ^2 $\mathcal{X}_n^2\left(\varphi^2\right)$ and t_* follows the non-central t distribution with $n-1$ degrees of freedom and non-centrality parameter δ, $t_{n-1}(\delta)$ with

$$\varphi^2 = \sum_{i=1}^{n} \left(\frac{\mu_i}{\sigma_i}\right)^2 \geq 0 \ \ and\ \ \delta^2 = \frac{n\mu^2}{\sigma^2}\ , \ \ with\ \ \varphi^2 \geq 0,\ \delta \in \mathbb{R}.$$

Their probability density function (pdf) of them is depending on the pdf of the central chi-square (f_{X^2}) and $t, (f_t)$ respectively and equal to

$$f_{X_*^2}(\omega) = f_{X^2}(\omega) \cdot exp\left\{-\frac{\varphi^2}{2}\right\} h_1\left(\frac{n}{2}, \frac{\varphi^2\omega}{4}\right),$$

$$f_{t_*}(\omega) = f_t(\omega) \cdot exp\left\{-\frac{\delta^2}{2}\right\} h_{\frac{1}{2}}\left(n+1, \frac{\delta\omega\sqrt{2}}{\sqrt{n^2+\omega^2}}\right).$$

Example 1.2 Consider the ratio on a non-central and central X^2 as

$$F_* = \frac{X_n^2\left(\varphi^2\right)/n}{X_m^2/m}.$$

Then F_* follows the non-central $F_{n,m}\left(\varphi^2\right)$ with pdf

$$f_{F_*}(\omega) = f_F(\omega) \cdot exp\left\{-\frac{\varphi^2}{2}\right\} H_{1,1}\left(\frac{m+n}{2}, \frac{m}{2}; \frac{\varphi^2 m\omega}{2(m+n)\omega}\right), \; with\; \varphi^2 \geq 0,\; \delta \in \mathbb{R}.$$

In this paper the emphasis is to the hypergeometric functions (HF) and not the sequential statistics. Therefore we are using statistical results for the main techniques, to prove, and provicle evidence that the HF play a dominant role to statistical applications, and not only to differential equations, were usually are adopted.

Consider the family of weighted functions

$$W = (w_i) \equiv \left\{1, (1-x)^{p+1}(1+x)^{q+1}, e^{-x}, x^{p+1}e^{-x}, e^{-x^2}\right\},$$

with $W \subseteq [-1,1]^2 \times [0,\infty)^2 \times \mathbb{R}$ and $p > -1,\; q > -1$. Then the following Theorem holds :

Proposition 1.1 *Consider the n-th degree polynomial of the form*

$$P_n(x) = \sum_{i=1}^{n} \beta_i x^{i-1},$$

and a weighting function from W. Then D-optimal design, admits equal weight $\frac{1}{n}$ at the (unique) roots of particular orthogonal polynomial special cases of Hypergeometric Functions, which forms the "optimal design points"

Proof. Due to Fedorov (1972, pg.88, [5]) for the corresponding weighting function the roots of the polynomials

i. $\left(1-x^2\right)\frac{dLeg_{n-1}(x)}{dx}$, with $Leg_n(x)$ the n-th degree Legendre polynomial.

ii. $J_n(x; a, \beta)$, with J_n the Jacobi polynomial with parameters a, β.

iii. $xLag_{n-1}(x;1)$

iv. $Lag_n(x;a)$ with Lag_n the Laguerre polynomial with parameter a.

v. $H_n(x)$, the n-th degree Hermite polynomial correspond to optimal design points.

But, consider the following relations, for HF in relation to the above:

i. $\frac{d}{dx}H_{1,1}(a,\beta;x) = \frac{a}{\beta}H_{1,1}(a+1,\gamma+1;x)$ and
$Leg_n(x) = \frac{1}{n!}(2x)^n \left(\frac{1}{2}\right)_n H_{2,1}\left(\frac{1}{2}-\frac{1}{2}n, -\frac{1}{2}n; \frac{1}{2}-n; \frac{1}{x^2}\right)$

ii. $J_n(x;a,\beta) = H_{2,1}(-n, a+n, \beta; x)$.

For iii and iv consider that $Lag_n(x;a) = \frac{(a+1)n}{n!}H_{1,1}(-n,a+1;x)$, and finally

v. $H_{2,n}(x) = (-1)^n \frac{(2n+1)!}{n!} 2zH_{1,1}\left(-n, \frac{3}{2}; x^2\right),\ x \geq 0.$

Thus the stated proposition has been proved. □

This Proposition 1.1 is an example of how we are facing statistical results: from the optical angle of the HF.

In the sequence the Sequential Probability Ratio Test (SPRT) is discuised due to which the presented results are linked with the Hypergeometric Functions.

2. The Sequential Probability Ratio Test (SPRT)

The main difference between the Classical and Sequential approach to Statistics is that the sample size in sequential approach is a random variable, while in Classical approach the sample size n is fixed. Moveover if the random variables $X_1, X_2, \ldots, X_n, \ldots$ are coming from the pdf $f(x;\vartheta)$ the test

$$H_0:\ \vartheta \in \Theta_0\ \ vs\ \ H_1:\ \vartheta \in \Theta_1\ \ with\ \ \Theta_0 \cap \Theta_1,\ \Theta_0,\ \Theta_1 \subseteq \Theta,$$

in Sequential Analysis is based in principle to the three choices rule:

Accept H_0 if $X_n \in A \subseteq \mathbb{R}$,
Regect H_0 if $X_n \in R \subseteq \mathbb{R}$,
Continue sampling if $X_n \in C \subseteq \mathbb{R}$,

with the acceptance (A), regection (R) and continuation (C) region to form a particion of R, while in Classical Statistics there is no C region

Subsect to $P\left(n < \infty|\vartheta\right) = 1$ which eventually means that the test is "closed" the Sequential Probality Ratio Test (SPRT), see Schervish (1995, section 9.2, [22]), is defined in contrast to the well known fixed one as

$$\Psi_n = \ln\frac{f\left(x_2, \vartheta_0\right)}{f\left(x_1, \vartheta_1\right)}\ ,\ \ n \geq 1.$$

The stopping boulds, and the continuation region $C = (L,U) \subseteq \mathbb{R}$ is defined as

$$\Psi_n = \begin{cases} \leq L,\ \ accept\ \ H_0, \\ \geq U,\ \ regect\ \ H_0, \\ \in (L,U),\ \ continue, \\ 0\ \ iff\ \ f(x;\vartheta_0) = 0 = f(x;\vartheta_1). \end{cases} \tag{1}$$

Due to the pionearing paper of Wald (1945) the optimal approximations of the lower and upper bounds are

$$L^* = \ln \frac{\beta}{1-\alpha},\ U^* = \ln \frac{1-\beta}{\alpha},$$

with, α, β being the type I and II errors, as usually. If we let

$$\Psi_i = \begin{cases} \ln \frac{f(x_i;\vartheta_0) f(x_{i-1};\vartheta_0)}{f(x_i;\vartheta_1) f(x_{i-1};\vartheta_1)}, & i > 1, \\ \ln \frac{f(x_1;\vartheta_0)}{f(x_1;\vartheta_1)}, & i = 1. \end{cases}$$

then, trivially $\Psi_n = \sum_{i=1}^{n} \Psi_i$.

Example 2.1 i. If $f(x;\vartheta)$ is the pdf of the normal $N(\mu, \sigma^2)$ distribution then

$$\Psi_i = \frac{x_i(\mu_i - \mu_0)}{\sigma^2} - \frac{(\mu_i - \mu_0)^2}{2\sigma^2}.$$

ii. If $f(x, \vartheta)$ is the pdf of the Exponential, $E_x(\vartheta)$, distribution, then

$$\Psi_i = \ln \frac{\vartheta_1}{\vartheta_0} - (\vartheta_1 - \vartheta_0) x_i.$$

Then at is was mentioned we continue iff

$$L^* < \Psi_n < U^*,$$

which is reduced to for the normal distribution to

$$\Lambda(L^*) + \lambda n \leq X_n \leq \Lambda(U^*) + \lambda n,$$

with

$$\Lambda(z) = \frac{z\sigma^2}{\mu_1 - \mu_0},\ \lambda = \frac{\mu_0 + \mu_1}{2},\ X_n = \Sigma_{i=1}^{n} X_i,$$

and for the exponential distribution to

$$\frac{1}{\vartheta_1 - \vartheta_0}\left[U^* + \ln\left(\frac{\vartheta_1}{\vartheta_0}\right)^n\right] \leq X_n \leq \frac{1}{\vartheta_1 - \vartheta_0}\left[L^* + \ln\left(\frac{\vartheta_1}{\vartheta_0}\right)^n\right],$$

or equivallently to

$$\Lambda(U^*) + \lambda n \leq X_n \leq \Lambda(L^*) - \lambda n,$$

with $\lambda = \frac{\ln \vartheta_1 - \ln \vartheta_0}{\vartheta_1 - \vartheta_0}$, $\Lambda(z) = \frac{z}{\vartheta_1 - \vartheta_0}$.

The Sequential Analysis is also related to the Stochastic Process, see Kitsos (2007) [13]. The multivariate approach due to Sequential approach has been discussed for Hotteling's T^2, Rencher (2002) [20], Kitsos (1993) [11], while the method appears an aesthetic appeal to optimal experimental theory, Kitsos (1989, 1992) [9],[10], and resently we have applied the Sequential principle to more realistic fields of applications, Kitsos and Hatzikian (2005) [12].

3. The SPRT and HF

In this section we shall discuss the relationship between the SPRT and the HF. The main results are from various statistical tests, Ghosh (1970) [7].

Proposition 3.1 *Let $X_1, X_2, \ldots, X_n, X_{n+1}, \ldots$ random variables, sequentially collected. If we define*

$$\overline{X_n} = n^{-1} \sum_{i=1}^{n} x_i \ , \quad S_n^2 = n^{-1} \sum_{i=1}^{n} (x_i - \overline{x_n})^2 \, .$$

Then

i. $\overline{x}_{n+1} = \frac{n}{n+1}\overline{x}_n + \frac{1}{n+1}x_{n+1}$

ii. $S_{n+1}^2 = \frac{n}{n+1}S_n^2 + \frac{n}{(n+1)^2}\left(\overline{X}_n - \overline{X}_{n+1}\right)^2$

ii. Let $U_n = \frac{\sum x_i}{\left(\sum x_i^2\right)^{\frac{1}{2}}}$. *Then*

$$U_n = \frac{n\overline{x}_n}{\left[n\left(\overline{x}_n^2 + S_n^2\right)\right]^{\frac{1}{2}}}. \tag{2}$$

Proof. Trivial. □

Although the proof is trivial it is essential to clarify the background: the mean and variance are evaluated iteratively. The quantity U_n is applied to the following:

Theorem 3.1 *(Sequential t-test) Given that $X_1, X_2, \ldots, X_n$ are independent identically distributed random variables from $N(\mu, \sigma^2)$ then the SPRT for testing*

$$H_0 \ : \ \frac{\mu}{\sigma} = \lambda_0 \quad vs \quad H_1 \ : \ \frac{\mu}{\sigma} = \lambda_1 \ , \ \lambda_0 < \lambda_1. \tag{3}$$

is a function of the confluent hypergeometric function.

Proof. Indeed: it is, see Ghosh (1970), for this particular case

$$\Psi_n = -\frac{n}{2}\left(\lambda_1^2 - \lambda_0^2\right) + \ln\frac{K(n,\gamma;\xi_1)}{K(n,\gamma;\xi_0)}, \tag{4}$$

with

$$K(n,\gamma;\xi_i) = H_{1,1}\left(\frac{n}{2}\ , \frac{1}{2}; \frac{1}{2}\xi_i^2\right) + \sqrt{2}\xi_i\gamma H_{1,1}\left(\frac{n+1}{2}\ , \frac{3}{2}; \frac{1}{2}\xi_i^2\right),$$

and $\xi_i = \lambda_i^2 U_n^2$, U_n^2 as in (iii) Proposition 3.1, $\gamma = \frac{\Gamma\left(\frac{1}{2}(n+1)\right)}{\Gamma\left(\frac{1}{2}n\right)}$. □

The above Theorem is extending with the following, see (3) and (5).

Theorem 3.2 *(Sequential F-tests) Given* $(X_{1i}, X_{2i}, \ldots, X_{pi})$, $i = 1, 2, \ldots, n$ *independent identically distributed p-dimensional random vectors, sutch that* $X_{ji} \sim N\left(\mu_j, \sigma^2\right)$, $j = 1, 2, \ldots, p$, $i = 1, 2, \ldots,$ *then the SPRT for testing*

$$H_0 : \frac{1}{\sigma^2} \sum_{j=1}^{p} \mu_j^2 = k_0 \quad vs \quad H_1 : \frac{1}{\sigma^2} \sum_{j=1}^{p} \mu_j^2 = k_1, \tag{5}$$

is a function of the confluent hypergeometric function.

Proof. Indeed, the SPRT is, see Ghosh (1970) [7]

$$\Psi_n = -\frac{n}{2}(k_1 - k_0) + \ln\left[\frac{A(k_1, V_n)}{A(k_0, V_n)}\right], \quad n \geq 1, \tag{6}$$

with

$$A(k_i, V_n) = H_{1,1}\left(\frac{pn}{2}, \frac{p}{2}; \frac{nk_iV_n}{2(V_n + 1)}\right), \ i = 0, 1,$$

and

$$V_n = \frac{\sum_{j=1}^{p} \omega_{ji}^2}{\sum_{j=1}^{p}\sum_{i=1}^{n} \omega_{ji}^2}\ , \quad \omega_{ji} = \frac{1}{\sqrt{n}} U_{ij} \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p.$$

□

Consider now the multivariate case such as the X_{ij} $j = 1, 2, \ldots, n$, $i = 1, 2, \ldots, n, \ldots$ observation can not necessarily be intependent with the same variance. Notice that i is not running as $i = 1, 2, \ldots, n$ as for the static design, but as the sequential approach is adopted we obtain $1, 2, \ldots, n, n+1, \ldots$ observation, for each vector x_j. The sample size is a random variable for the SPRT.

The random vector $(X_1, X_2, \ldots, X_k)$ we say that follows the multivariate normal $N_k(\mu, \sum)$ if the joint probability density functions (pdf) of $(x_1, x_2, \ldots, x_k)$ is of the form

$$f\left(X; \mu, \Sigma\right) = (2\pi)^{-\frac{1}{2}} \det \Sigma^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} Q\left(\mu, \Sigma\right)\right\},$$

$$Q\left(\mu, \Sigma\right) = \langle (x - \mu), \Sigma^{-1}(x - \mu)^t \rangle,$$

where $\langle a, b \rangle$ is the inner product of the vector a,b and b^t is the transpose of b.

Theorem 3.3 *(Sequential* T^2*-test). The SPRT for testing*

$$H_0 : \ \mu\Sigma^{-1}\mu^t \leq r_0 \quad vs \quad \mu\Sigma^{-1}\mu^t \geq r_1, \ \ 0 \leq r_0 < r_1,$$

is a function of HF.

Proof. Indeed, the SPRT is, see Ghosh (1970) [7]

$$\Psi_n = -\frac{n}{2}(r_1 - r_0) + \ln\frac{M(n,k,r_1,V_n)}{M(n,k,r_0,V_n)}, \tag{7}$$

with $M(n,k,r_1,V_n) = H_{1,1}\left(\frac{n}{2}, \frac{k}{2}; \frac{nr_iV_n}{(1+V_n)}\right)$ and

$$V_n = \frac{1}{\sigma^2}\sum_{j=1}^{n}\bar{x}^2_{j(n)}\ , \quad \bar{x}^2_{j(n)} = n^{-1}\sum_{i=1}^{n}x_{ji}.$$

□

From the above stated Theorems, from the point of view of HF it is clear that HF are essential to the main SPRT.

In the next section the problem of sequential regression is discussed through the SPRT for the correlation coeficient, Ghosh (1970) [7], while the classical approach is disussed in the pioneering work of Anderson (1958) [1], Graybill (1976) [8] among others.

4. Sequential Regression and HF

Now let us consider the pairs of observations $z_n = (x_i, y_i),\ i = 1, 2, \ldots, n$ coming from the bivariate normal distribution of the form, with $p \in (-1, 1)$ being the correlation coefficient

$$f_n(z_n;\vartheta) = \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^{-n} exp\left\{-\frac{1}{2(1-p^2)}Q_n(\vartheta)\right\},$$

with $\vartheta = (\mu_1, \mu_2, \sigma_1, \sigma_2) \in \mathbb{R}^2 \times \mathbb{R}^2_+$ and

$$Q_n(\vartheta) = \sum_{j=1}^{n} z^2_{ij} - 2\rho\sum_{i=1}^{n} z_{i1}z_{i2} + \sum_{i=1}^{n} z^2_{i2},$$

and

$$z_{ij} = \frac{x_i - \mu_i}{\sigma_j}\ ,\ j = 1, 2, \quad i = 1, 2, \ldots, n.$$

Lemma 4.1 *For $n > 2$ the pdf of the estimate $r = r_n$ of ρ, $f_n(r;\rho)$ is a function of the hypergeometric function $H_{2,1}$.*

Proof. Indeed, see Anderson (1958) [1], Graybill (1976) [8]

$$f_n(r;\rho) = \gamma_n a^p_{rr} a^q_{\rho\rho} a^s_{\rho r} H_{2,1}\left(\frac{1}{2}\ , \frac{1}{2}\ , n - \frac{1}{2}; \frac{1 + r\rho}{2}\right),$$

with

$$\gamma_n = \frac{n-2}{\sqrt{2\pi}}\frac{\Gamma(n-1)}{\Gamma\left(n-\frac{1}{2}\right)}\ , \quad a_{\varrho r} = 1 - \rho r\ , p = \frac{n-4}{2}\ , \quad q = \frac{n-1}{2}\ , \quad s = -n + \frac{3}{2}.$$

□

Notice that $-1 < r,\ \rho < 1$, so the defined $H_{2,1}$ exists. Based on the above Lemma the follow theorem can be proved.

Theorem 4.1 *Given the paired observations* $z_n = (x_i, y_i),\ i = 1, 2, \ldots, n, n+1, \ldots$, *coming from the bivariate normal distribution, the SPRT for testing*

$$H_0 : \rho = \rho_0 \quad vs \quad H_1 : \rho = \rho_1\ , -1 < \rho_0 < \rho_1 < 1,$$

is a function of the HF.

Proof. Indeed, Ghosh (1970) [7],

$$\Psi_n = \frac{n-1}{2} \ln \frac{a_{\rho_1\rho_1}}{a_{\rho_0\rho_0}} - \left(n - \frac{3}{2}\right) \ln \frac{a_{\rho_1 r_n}}{a_{\rho_0 r_n}} + \ln \frac{A(\rho_1, r_n)}{A(\rho_0, r_n)}, \tag{8}$$

with

$$A(\rho_i, r_n) = H_{2,1}\left(\frac{1}{2}\ , \frac{1}{2}\ , n - \frac{1}{2}; \frac{1+\rho_i r_n}{2}\right),\ i = 1, 0,\ n > 2,$$

and $a_{\lambda\mu} = 1 - \lambda\mu$. □

The following Corollary can be proved useful to the practicians as they can know when a performed simple regression is valid, before performing it. This Corollary 4.1 is essential to economical problems, were for a sort period of time the linear regression is assumed, without being clear if the regression is valid.

Corollary 4.1 *The regression equation* $y_i = \beta_o + \beta_1 x + e_i$, *and the estimation* $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ *is valid only when*

$$\Psi_n \leq \ln \frac{\beta}{1-a},$$

for the defined ρ_0, *and assuming* $e_i \sim N(0, \sigma^2)$.

Proof. Due to Theorem 4.1, relation (8) and the definition of Ψ_n, relation (1) in section 2. □

In a physical problem it can be $0.90 < \rho_0 < 0.95$, i.e. the correlation is high, while for economical or social problems can be accepted even around 0.70, i.e. lower correlation. Example: For $a = 0.01$ and while $\beta = 0.05$, $\Psi_n \leq -2.985$.

Corollary 4.2 *When*

$$\Psi_n \geq \ln \frac{1-\beta}{a},$$

resect H_0. *This is equivalent to*

$$\left| \frac{\arctan(\rho_0) - \arctan(r)}{\sqrt{n-3}} \right| \geq Z_{\frac{a}{2}}.$$

Proof. Due to definition of Ψ_n in (8) and Theorem 11.2.5 in Graybill (1976) [8] as we are rejects H_0, and $Z_{\frac{a}{2}}$ as usually. □

Notice that the power evaluated by Graybill (1976) [8] can not be applied in the sequential case, as there are three regions $(A, \mathbb{R}, \mathbb{C})$ under consideration.

5. Approximations to SPRT

The crucial problem is that the Statistician (and not only) has a problem to evaluate the SPRT Ψ_n as have been proposed to the Theorems in sections 3 and 4. Thus, approximations are need to simplify for the statistician the above relations, see (4), (6), (7) and (8). The main idea is that U_n can be either $U_n^2 \sim n$ or $U_n^2 \sim n^2$ to calculate a number of points (n, Ψ_n) and to apply the regression method to fit a linear model. This method can be used to other SPRT as well.

Therefore, we propose the following approximations for (4), easy to be applied.

Proposition 5.1 *When $U_n^2 = 4n$ then the linear approximation of (4) holds:*

$$\Psi_n^* = \begin{cases} 29.785n - 1.1578, \ \textit{with}\ \lambda_0 = 1,\ \lambda_1 = 2, \\ 2.3541n - 0.4962, \ \textit{with}\ \lambda_0 = 0.5,\ \lambda_1 = 1, \\ 0.0466n - 0.0148, \ \textit{with}\ \lambda_0 = 0.1,\ \lambda_1 = 0.2. \end{cases} \tag{9}$$

Proposition 5.2 *When $U_n^2 = 4n^2$ then the quadratic approximation of (4) holds:*

$$\tilde{\Psi}_n = \begin{cases} 29.9974n^2 - 0.1281n - 1.3466, \ \textit{with}\ \lambda_0 = 1,\ \lambda_1 = 2, \\ 1.8868n^2 + 0.7366n - 0.9164, \ \textit{with}\ \lambda_0 = 0.5,\ \lambda_1 = 1, \\ 0.0091n^2 + 0.1080n - 0.1389, \ \textit{with}\ \lambda_0 = 0.1,\ \lambda_1 = 0.2. \end{cases} \tag{10}$$

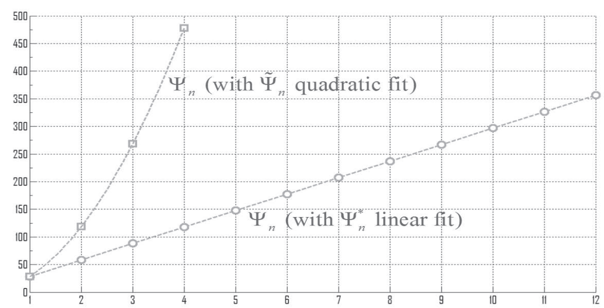

Figure 1: Graphs of the actual Ψ_n values (with $\lambda_0 = 1$, $\lambda_1 = 2$) and their approx. linear regression fit Ψ_n^* and $\tilde{\Psi}_n$.

For the above proposed approximations (9) and (10) the corresponding $R^2 \cong 1$ and therefore there are statistical evidence that the model is valid. Figure 1 clarifies the proposed approximation.

Proposition 5.3 *For $\lambda_0 = 1, \lambda_0 = 2, U_n^2 = 4n$ the SPRT for the sequential t-test is approximated to the iterative scheme*

$$\Psi_n^{**} = 3.090 \times 6^n - 2.090 \times (-5)^n. \tag{11}$$

Proof. For the given values it is calculated $(1, 28.5230)$, $(2, 58.3860)$, $(3, 88.1977)$, $(4, 117.8)$ and therefore the iterative scheme

$$\Psi_n^{**} = \Psi_{n-1}^{**} + 30,$$

approximates these values. The characteristic equation of the above different equation provides roots $\varphi_1 = 6,\ \varphi_2 = -5$ and due initial values Ψ_1, Ψ_2, (11) is true. □

The above result is rather useful as for dimension $n > 4$ the particular HF is not easily calculated with $a = 0.05,\ \beta = 0.10$, see (1). The idea behind is provide the values you wish for your a priori experience and approximate the problem with a difference equation.

The same way of thinking is adopted for the SPRT, see (6) in Theorem 3.2.

Proposition 5.4 *When $V_n^2 = 9n$ then (6) is approximated as*

$$\Psi_n^* = \begin{cases} A_p n + B_p, \\ a_p n^2 + b_p n + c_p, \end{cases}$$

with $k_0 = 1,\ k_1 = 2$, where

$$(A_p, B_p) = \begin{cases} (0.3734, -0.4914),\ p = 2\ (R^2 \cong 1), \\ (0.4941, -0.6844),\ p = 3\ (R^2 \cong 1), \\ (0.6888, -1.0358),\ p = 5\ (R^2 \cong 1), \end{cases}$$

and

$$(a_p, b_p, c_p) = \begin{cases} (2 \cdot 10^{-3}, 0.3658, -0.4510),\ p = 2\ (R^2 \cong 1), \\ (4 \cdot 10^{-3}, -0.4822, -0.6210),\ p = 3\ (R^2 \cong 1), \\ (7 \cdot 10^{-3}, -0.6661, -0.9149),\ p = 5\ (R^2 \cong 1). \end{cases}$$

Proposition 5.5 *When $V_n^2 = 9n^2$ then (6) is approximated as*

$$\tilde{\Psi}_n = \begin{cases} K_p n + L_p, \\ k_p n^2 + l_p n + m_p, \end{cases}$$

with $k_0 = 1,\ k_1 = 2$, where

$$(K_p, L_p) = \begin{cases} (0.3877, -0.4968),\ p = 2\ (R^2 \cong 1), \\ (0.5096, -0.6361),\ p = 3\ (R^2 \cong 1), \\ (0.7065, -0.9818),\ p = 5\ (R^2 \cong 1), \end{cases}$$

and

$$(k_p, l_p, m_p) = \begin{cases} (4 \cdot 10^{-5}, 0.3891, -0.4541),\ p = 2\ (R^2 \cong 1), \\ (10^{-4}, -0.5075, -0.6248),\ p = 3\ (R^2 \cong 1), \\ (4 \cdot 10^{-4}, -0.6950, -0.9204),\ p = 5\ (R^2 \cong 1). \end{cases}$$

Figure 2 clarifies the proposed approximation.

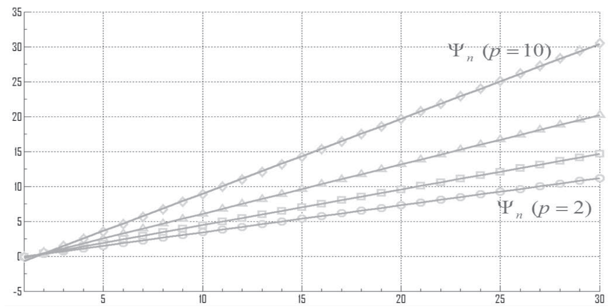

Figure 2: Graphs of the actual Ψ_n values (with $k_0 = 1$, $k_1 = 2$) and their approx. linear regression fit $\tilde{\Psi}_n$.

Proposition 5.6 *Fot the SPRT as in (8) the piece-wise regression model provides an approximation for it, i.e.*

$$\tilde{\Psi}_n = \begin{cases} 0.0099n^2 - 0.4934n + 0.5130, \ for \ 3 \le n \le 23 \ (R^2 = 0.9998), \\ 0.0232n^2 - 1.1353n + 8.35, \ for \ 23 \le n \le 40 \ (R^2 = 0.9993), \\ 0.0557n^2 - 3.7302n + 60.2144, \ for \ 40 \le n \le 50 \ (R^2 = 0.9999). \end{cases}$$

for $\rho_0 = 0.6$, $\rho_1 = 0.9$.

Figure 3 clarifies the above proposed approximation.

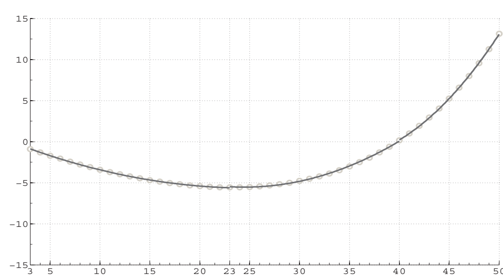

Figure 3: Graphs of the actual Ψ_n values (with $\rho_0 = 0.6$, $\rho_1 = 0.9$) and their approx. piece-wise quadratic regression fit $\tilde{\Psi}_n$.

6. Discussion

The HF are mainly applied to the solution of differential equations, with the Schrödinger's equation for the ration of a symmetrical-tope molecule, being a typical example. There is also a generalization from $H_{1,1}, H_{2,1}$ to $H_{p,q}$ target was to link this excellent mathematical tool with a class of Statistical applications. We tried to prove that really the application of HF can be succesfully extended to Statistics with typical situation being the optimal experimental design point. The SPRT for a number of tests can be easily approximated. The whole set of approximations we have, is still extended. Moreover the complicated, for a practicia Statistician, form of the SPRT have been approximated with simple forms; roughly speaking when $U_n \sim n$, Ψ_n is linear form and when $U_n \sim n^2$ is quadratic form can be adopted. Proposition 5.6 provides the approximation with the sample size to play a dominant role. Therefore, we believe that HF plays an important role, but in practice, statisticians can adopt the proposed approximations for their calculations.

I would like to thank Dr. Alexandros Pappas for usefull discussions and Dr. Thomas Toulias for the support in computations.

References

1. Anderson, T. W., *An Introduction to Multivariate Statistical Analysis.* John Willey, New York, 1958.
2. Anderson, T. W., *A Modification of the Sequential Probality Ratio Test to Reduce the Sample Size.* Ann. Math. Stat., Vol. 31, No. 1 (1960), 165-197.
3. Arfken, G. and Weber, H., *Mathematical Methods for Physicists.* Academic Press, 2000.
4. Bayin, S. S., *The impact of loan Methods in Science and Engineering.* Wiley, 2006.
5. Fedorov, V. V., *Theory of Optimal Experiments.* Academic Press, 1972.
6. Gautschi, W., *Gauss quadrature approximates to hypergeometric and confluent hypergeometric functions.* J. of Comp. and Appl. Math., Vol. 139 (2002), 173-187.
7. Ghosh, B. K., *Tests of Statistical Hypotheses.* Addison-Wesley Pub. Co. 1970.
8. Graybill, F., *Fully sequential procedures in Nonlinear Design Problems.* Theory and Application of the Linear Model, Duxbury Press, 1976.
9. Kitsos, C. P., *Fully sequential procedures in Nonlinear Design Problems.* Comp. Stat. and Data Analysis, Vol. 8 (1989), 13-19.
10. Kitsos, C. P., *Quasi-Sequential Procedures for the Calibration Problem.* MODA 3 (1992), 227-231.
11. Kitsos, C. P., *Multiple - Multivariate - Sequential T^2 comparisons.* LINSTAT' 93, Proceedings, T. Calinski, R. Kala (Eds), 47-51.
12. Kitsos, C. P. and Hatjikian, J., *Sequential Techniques for Innovation Indexes.* 16ISPIM, e-volume, 2005.
13. Kitsos, C. P., *The SPRT for the Poisson.* ASMDA Chania, May 29 - June 1, 2007.
14. Kitsos, C. P. and Tavoularis, N., *Logarithm Sobolev Inequalities for information Measures.* IEEE Trans Inform. Theory, Vol. 55, No. 6 (2009), 2554-2561.
15. Lebedev, N. N., *Special functions and their Applications.* Prentice-Hall, 1956.
16. Morrison, D. F., *Multivariate Statistical Methods.* Mc Graw-Hill, New York, 1976.
17. Owen, D. B., *A survey of Proprties and Applications of the non-central t-Distribution.* Techometrics, Vol. 10 (1968), 445-468.

18. Antsaklis, P. J., and Michel A. N., *Linear systems.* The McGraw-Hill Companies, INC. U.S.A., 1997.
19. Pukelsheim, F., *Optimal Design of Experiments.* John Wiley, Toronto, 1993.
20. Rencher, A. C., *Methods of Multivariate Analysis.* Wiley-Interscience, 2002.
21. Sanchez-Ruiz, J. and Dehesa, J. S., *Fisher information of orthogonal hypergeometric polynomials.* J. of Comp. and App. Math., Vol. 182 (2005), 150-164.
22. Scherrish, J. M., *Theory of Statistics.* Springer-Verlag, New York, 1995
23. Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry.* Oliver and Boyd, Edinburgh, 1966.
24. Wald, A., *Sequential tests of statistical Hypotheses.* Ann. Math. Stat., Vol. 16 (1945), 117-186.

◇ C. P. Kitsos
Department of Mathematics
Technological Educational Institute (T.E.I) of Athens
GR-122 10, Egaloeo, Athens, GREECE
xkitsos@teiath.gr