

11-1-2011

Comparison of Several Tests for Combining Several Independent Tests

Madhusudan Bhandary

Columbus State University, bhandary_madhusudan@columbusstate.edu

Xuan Zhang

Pracs Institute, Ltd., zhangxuannd@yahoo.com

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>



Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Bhandary, Madhusudan and Zhang, Xuan (2011) "Comparison of Several Tests for Combining Several Independent Tests," *Journal of Modern Applied Statistical Methods*: Vol. 10 : Iss. 2 , Article 5.

DOI: 10.22237/jmasm/1320120240

Available at: <http://digitalcommons.wayne.edu/jmasm/vol10/iss2/5>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Comparison of Several Tests for Combining Several Independent Tests

Madhusudan Bhandary
Columbus State University
Columbus, GA

Xuan Zhang
Pracs Institute, Ltd.
Fargo, ND

Several tests for combining p-values from independent tests have been considered to address a particular common testing problem. A simulation study shows that Fisher's (1932) Inverse Chi-square test is optimal based on a power comparison of several different tests.

Key words: Omnibus test, omnibus hypothesis, p-value, Kolmogorov-Smirnov test, Tippett's test, Wilkinson's test, Inverse Chi-square test, Inverse normal test, Logit test.

Introduction

Tests for statistical significance of combined results were possibly the first statistical procedures developed for quantitative research synthesis. Combined test procedures were developed to combine the results of significance tests from different research studies.

Combining data from similar studies, as opposed to data derived from a single study, is important in Statistics. This study is a review of so-called omnibus statistical methods for testing the statistical significance of combined results. The procedures are called omnibus or non-parametric because they do not depend on the form of the underlying data, but only on the exact significance levels commonly called p-values. A key point is that observed p-values derived from continuous test statistics have a uniform distribution under the null hypothesis regardless of the test statistics or distribution from which they arise. The non-parametric nature of combined significance tests gives great flexibility in applications. Such tests can be used to combine any independent tests of hypotheses, even though the individual tests examine somewhat different hypotheses. For example,

combined significance tests may be used to summarize the results of 10 studies each of which examined the effect of a treatment on a different outcome variable. Such a procedure would test whether the treatment produced a superior outcome on any of the dimensions investigated. These procedures can also be used in research synthesis to combine the results of studies that test the same conceptual hypothesis by different methods.

Many statistical tests are available for testing the significance for combining results. This study examines the most widely used tests. Nine different tests were compared, these are: Kolmogorov-Smirnov, Tippett's, Wilkinson's (for $r = 2, 3, 4, 5$), Inverse Chi-square, Inverse normal and Logit test. The objective of this study was to perform a comprehensive comparison of the performance of these tests based on their powers. A simulation study was conducted and the powers of the tests were compared. It was found that Fisher's (1932) Inverse Chi-square test was optimal based on the power comparison of the different tests.

p-Value Calculation: Normal Distribution

Let X_1, X_2, \dots, X_n be a random sample from $N(\mu, \sigma^2)$. Let \bar{X} be the sample mean and let u be the observed value of the sample mean. Let $\Phi(\cdot)$ be the distribution function of the standard normal distribution.

Madhusudan Bhandary is a Professor in the Department of Mathematics. Email him at: bhandary_madhusudan@columbusstate.edu.
Xuan Zhang is a statistician. Email him at: zhangxuannd@yahoo.com.

Case 1

$$H_0 : \mu = \mu_0(\text{specified})$$

$$H_1 : \mu < \mu_0$$

$$\Lambda = \frac{\text{Sup}_{H_0} L(\mu)}{\text{Sup}_{\Omega} L(\mu)} = \left(\frac{\bar{X}}{\mu_0}\right)^n e^{-\frac{n\bar{X}}{\mu_0} + n}$$

$$p\text{-value} = \Pr(\bar{X} \leq u / H_0)$$

$$= P\left(\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \leq \frac{\sqrt{n}(u - \mu_0)}{\sigma}\right)$$

$$= \Phi\left(\frac{\sqrt{n}(u - \mu_0)}{\sigma}\right)$$

where, Ω is the parameter space. For large n,

$$p\text{-value} = P(\Lambda \leq \text{observed } \Lambda / H_0)$$

$$= P(-2 \ln \Lambda \leq -2 \ln(\text{observed } \Lambda) / H_0)$$

$$= P(\chi_1^2 \geq -2 \ln(\text{observed } \Lambda))$$

Case 2

$$H_0 : \mu = \mu_0(\text{specified})$$

$$H_1 : \mu > \mu_0$$

$$p\text{-value} = \Pr(\bar{X} \geq u / H_0)$$

$$= 1 - P(\bar{X} < u / H_0)$$

$$= 1 - P\left(\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} < \frac{\sqrt{n}(u - \mu_0)}{\sigma}\right)$$

$$= 1 - \Phi\left(\frac{\sqrt{n}(u - \mu_0)}{\sigma}\right)$$

Omnibus Hypotheses and Omnibus Tests

Suppose n independent investigators have set about testing the validity of some null hypothesis:

$$H_0 : \text{The population mean is } \mu_0(\text{specified})$$

versus

$$H_1 : \text{The population mean } \mu < \mu_0$$

Each investigator will select a random sample from the population under focus, collect the relevant data, apply the appropriate test, and then report the p -value. The sample size could vary from investigator to investigator. The information provided by the investigators can be summarized as follows:

Case 3

$$H_0 : \mu = \mu_0(\text{specified})$$

$$H_1 : \mu \neq \mu_0$$

$$p\text{-value} = P\left(\left|\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma}\right| \geq \left|\frac{\sqrt{n}(u - \mu_0)}{\sigma}\right| / H_0\right)$$

$$= 2 \left[1 - \Phi\left(\left|\frac{\sqrt{n}(u - \mu_0)}{\sigma}\right|\right) \right]$$

Investigator	Sample Size	p -Value
1	n_1	p_1
2	n_2	p_2
...
n	n_n	p_n

p-Value Calculation: Exponential Distribution

Let X_1, X_2, \dots, X_n be a random sample from $\text{EXP}(\mu)$. Let \bar{X} be their sample mean.

$$H_0 : \mu = \mu_0(\text{specified})$$

$$H_1 : \mu \neq \mu_0$$

The likelihood ratio test is given by:

The objective is to determine if the null hypothesis is universally true. If the null hypothesis is true overall then, theoretically, p_1, p_2, \dots, p_n should be a random sample of size n from a uniform distribution over $(0, 1)$. In order to test the merit of the hypothesis overall, a test statistic must be built that is a function of the data p_1, p_2, \dots, p_n . A multitude of tests have been proposed in this connection, but before presenting a plethora of tests, the above problem must be generalized.

COMPARISON OF TESTS FOR COMBINING SEVERAL INDEPENDENT TESTS

Assume m independent investigators, each investigating a hypothesis testing problem where H_{0i} is the null hypothesis proposed by investigator i , and H_{1i} is the alternative $i = 1, 2, \dots, m$. Each investigator collects data, tests his/her hypothesis and reports a p -value. This scenario can be summarized as follows:

Investigator	Null Hypothesis	Alternative Hypothesis	p -value
1	H_{01}	H_{11}	p_1
2	H_{02}	H_{12}	p_2
...
m	H_{0m}	H_{1m}	p_m

Postulating that the omnibus hypothesis, $H_0: H_{0i}$ is true for all i , versus the alternative $H_1: \text{at least one } H_{1i}$ is true, the data to decide in this issue are p_1, p_2, \dots, p_m . Theoretically, each p_i has a uniform distribution over $(0, 1)$ if H_{0i} is true. If the omnibus null hypothesis is true, p_1, p_2, \dots, p_m are independently, identically uniformly distributed over $(0, 1)$. Now replace both the omnibus null and alternate hypotheses with the following equivalent hypotheses:

$H_0: p_1, p_2, \dots, p_m$ is a random sample from a uniform distribution over $(0, 1)$,
versus

$H_1: p_1, p_2, \dots, p_m$ is a random sample from a distribution which is not a uniform distribution over $(0, 1)$.

Several tests have been developed to test the validity of the above modified hypotheses.

Test 1: Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (KS) test was originally proposed in the 1930's by Kolmogorov (1933) and Smirnov (1939). The KS test is only appropriate for testing data against a continuous distribution. The KS test statistic is defined as follows:

$$D = \sup_{0 < p < 1} |\hat{F}(p) - p|,$$

where \hat{F} is the empirical distribution function of the data p_1, p_2, \dots, p_m . The exact distribution of D under H_0 has been worked out and a table of critical values is available.

Test 2: Tippett's Test

The first test of the significance of combined results was proposed by Tippett (1931), who noted that, if p_1, p_2, \dots, p_m are independent p -values from continuous test statistics, then each has a uniform distribution under H_0 . The test procedure is as follows: Reject H_0 if $p_{(1)} < 1 - (1 - \alpha)^{1/m}$, where $p_{(1)} = \text{minimum of } p_1, p_2, \dots, p_m$. The p -value of the test is $= 1 - (1 - p_{(1)})^m$.

Test 3: Wilkinson's Test

Wilkinson (1951) provided a generalization of Tippett's procedure that uses not just the smallest but the r^{th} smallest p -value, $p_{(r)}$, as a test statistic, where $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)}$ are the ordered p -values (order statistics) obtained from p_1, p_2, \dots, p_m . The test procedure is given as follows: Reject H_0 if $p_{(r)} < p_{r,\alpha}$, where $p_{r,\alpha}$ is a critical value for $p_{(r)}$, or use a critical number $m_{r,\alpha}$ of p -values that are smaller than a fixed level α . Wilkinson described his procedure in terms of the number of significant p -values, that is, those that are smaller than α . He provided tables of the probability of obtaining m or more significant results at the $\alpha = 0.05$ and $\alpha = 0.01$ levels (that is, m or more p -values less than 0.05 or 0.01) for $m < 25$. Nomographs extending Wilkinson's tables to $m = 100$ for $\alpha = 0.05$ and to $m = 500$ for $\alpha = 0.01$ are given in Sakoda, Cohen and Beall (1954). Because $p_{(r)}$ has a beta distribution with parameters r and $n-r+1$, tables of the incomplete beta function can be used to obtain critical values of $p_{(r)}$ directly.

Test 4: The Inverse Chi-Square Test

One of the most widely used combination procedures is from Fisher (1932). Given m independent studies and p -values p_1, p_2, \dots, p_m , Fisher's procedure uses the product $p_1 p_2 \dots p_m$ to combine p -values. He used a connection between the uniform distribution and the Chi-square distribution – namely, that if u has a uniform distribution, then $-2 \ln u$ has a Chi-square distribution with 2 degrees of freedom. Consequently, when H_{0i} is true, $-2 \ln p_i$ has a Chi-square distribution with 2 degrees of freedom, therefore, $-2 \ln(p_1 p_2 \dots p_m) = \sum_{i=1}^m -2 \ln p_i$ also has a Chi-square distribution with $2m$ degrees of freedom. Due to this fact, no special tables are needed for the Fisher method. The test procedure becomes, reject H_0 if $T = -2 \sum_{i=1}^m \ln p_i \geq c$, where the critical value c is obtained from the upper tail of the chi-square distribution with $2m$ degrees of freedom.

Test 5: The Inverse Normal Test

Another procedure for combining p -values is the inverse normal method proposed independently by Stouffer, et al. (1949) and by Liptak (1958). This procedure involves transforming each p -value to the corresponding normal score and then averaging. More specifically, defining Z_i by $p_i = \Phi(Z_i)$, where $\Phi(x)$ is the standard normal cumulative distribution function. When H_0 is true, the statistic

$$Z = \frac{Z_1 + Z_2 + \dots + Z_m}{\sqrt{m}} = \frac{\Phi^{-1}(p_1) + \Phi^{-1}(p_2) + \dots + \Phi^{-1}(p_m)}{\sqrt{m}}$$

has the standard normal distribution. Thus, H_0 is rejected whenever Z exceeds the appropriate critical value of the standard normal distribution.

Test 6: The Logit Test

The method of combining m independent p -values, p_1, p_2, \dots, p_m , suggested by George (1977) and investigated by Mudholkar and George (1979) transforms each p -value into a logit, $\ln\left(\frac{p}{1-p}\right)$, and then combine the logits via the statistic

$$L = \ln \frac{p_1}{1-p_1} + \dots + \ln \frac{p_m}{1-p_m}.$$

The exact distribution of L is not simple, but when H_0 is true, Mudholkar and George (1979) showed that the distribution of L (except for a constant) can be closely approximated by Student's t -distribution with $5m+4$ degrees of freedom. Therefore, the test procedure is reject

H_0 if $L^* = |L| \sqrt{\frac{(0.3)(5m+4)}{m(5m+2)}} > c$ where the critical value c is obtained from the t -distribution with $5m+4$ degrees of freedom. (Note that the term 0.3 is more accurately given by $\frac{3}{\pi^2}$.) For

large values of m , $L^* \approx \left(\frac{0.55}{\sqrt{m}}\right) L$.

Methodology

Monte Carlo Simulation

A Monte Carlo simulation study was conducted to compare the performance of the omnibus test statistics described on the basis of estimated powers when the underlying data distributions are normal and exponential. The sample sizes used were 10 and 100. The omnibus hypotheses are:

$$H_0 : \mu = 5$$

versus

$$H_1 : \mu \neq 5$$

The maintenance of significance levels was checked for each of the nine tests (for Test 3, $r = 2, 3, 4, 5$ were used), under each sample size and population mean, and for two distributions:

COMPARISON OF TESTS FOR COMBINING SEVERAL INDEPENDENT TESTS

normal and exponential. Empirical error rates for each case were estimated by first simulating 10,000 different samples with specified sample size and population mean (μ_0) from a population with a specified distribution.

The test of interest was performed on each sample and it was determined if the null hypothesis was rejected at the 5% significance level. The empirical error rates for that test were then computed as the proportion of times the null hypothesis was rejected at each significance level. A test was considered acceptable at the 5% significance level if the error rates were between 0.044 and 0.056. The range represents a 99% confidence interval for the stated significance level.

Results

Tables 1-4 display the estimated powers of each test statistic investigated at the 0.05 significance level; Figures 1-4 show the power curves.

Conclusion

Of the nine test statistics considered, the Inverse Chi-square test gives the highest power in almost every simulation, regardless of the number of populations, sample size or parameter values. The second highest power observed was with the Inverse Normal test. The minimum p test almost always gave the lowest power. In general, the Inverse Chi-Square proved superior by performing consistently in simulations for a wide range of cases.

References

- Fisher, R. A. (1932). *Statistical methods for research workers* (4th Ed.). London: Oliver & Boyd.
- Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Institute of Italian Attuari. Gorn.*, 4, 1-11.
- Smirnov, N. V. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bulletin of the University of Moscow*, 2(2), 3-14.
- Tippett, L. H. C. (1931). *The methods of statistics* (1st Ed.). London: Williams & Norgate.
- Wikinson, B. (1951). A statistical consideration in psychological research. *Psychological Bulletin*, 48, 156-158.
- Sakoda, J. M., Cohen, B. H., & Beall, G. (1954). Test of significance for a series of statistical tests. *Psychological Bulletin*, 51, 172-175.
- Stouffer, S. A., Suchman, E. A., Devinney, L. C., Star, S. A., & Williams, R. M., Jr. (1949). *The American soldier 1949: Adjustment during army life*. Princeton, NJ: Princeton University Press.
- Liptak, T. (1958). On the combination of independent tests. *Magyar Tudományok Akademia Matematikai Kutató Intézetének Közleményei*, 3, 1971-1977.
- George, E. O. (1977). *Combining independent one-sided and two-sided statistical tests: Some theory and applications*. Unpublished doctoral dissertation, University of Rochester.
- Mudholkar, G. S., & George, E. O. (1979). The logit method for combining probabilities. In *Symposium on optimizing methods in statistics*, J. Rustagi (Ed.). 345-366. New York, NY: Academic Press.

Table 1: Normal Distribution $n = 10$, $nrep = 10,000$, $\alpha = 0.05$

μ	KS	P(1)	P(2)	P(3)	P(4)	P(5)	INV-CHI	INV-NORM	LOGIT
4.0	0.9712	0.8352	0.9467	0.9705	0.9730	0.9709	0.9926	0.9901	0.9827
4.1	0.9171	0.7380	0.8915	0.9241	0.9286	0.9259	0.9735	0.9657	0.9441
4.2	0.8108	0.6101	0.7879	0.8354	0.8443	0.8328	0.9225	0.9025	0.8556
4.3	0.6635	0.4830	0.6530	0.6973	0.7105	0.6910	0.8096	0.7796	0.7009
4.4	0.4901	0.3691	0.4919	0.5300	0.5412	0.5190	0.6511	0.6060	0.5119
4.5	0.3308	0.2643	0.3514	0.3738	0.3705	0.3607	0.4632	0.4254	0.3351
4.6	0.2051	0.1830	0.2253	0.2384	0.2351	0.2226	0.2982	0.2658	0.1891
4.7	0.1275	0.1158	0.1315	0.1418	0.1395	0.1296	0.1614	0.1488	0.1000
4.8	0.0817	0.0816	0.0853	0.0845	0.0844	0.0852	0.0946	0.0871	0.0658
4.9	0.0622	0.0578	0.0568	0.0589	0.0595	0.0552	0.0601	0.0586	0.0492
5.0	0.0529	0.0501	0.0509	0.0505	0.0517	0.0525	0.0492	0.0483	0.0472
5.1	0.0610	0.0576	0.0575	0.0580	0.0587	0.0573	0.0629	0.0595	0.0533
5.2	0.0835	0.0785	0.0822	0.0839	0.0856	0.0821	0.0913	0.0848	0.0622
5.3	0.1168	0.1194	0.1399	0.1410	0.1360	0.1263	0.1667	0.1460	0.0999
5.4	0.1975	0.1750	0.2214	0.2306	0.2301	0.2168	0.2887	0.2618	0.1860
5.5	0.3328	0.2599	0.3433	0.3743	0.3767	0.3598	0.4649	0.4187	0.3308
5.6	0.4853	0.3651	0.4985	0.5387	0.5352	0.5186	0.6506	0.6089	0.5176
5.7	0.6747	0.4825	0.6502	0.7005	0.7130	0.6946	0.8148	0.7789	0.7044
5.8	0.8079	0.6088	0.7831	0.8307	0.8363	0.8272	0.9166	0.8963	0.8484
5.9	0.9159	0.7273	0.8864	0.9264	0.9261	0.9197	0.9726	0.9631	0.9404
6.0	0.9688	0.8351	0.9518	0.9733	0.9746	0.9707	0.9935	0.9918	0.9832

COMPARISON OF TESTS FOR COMBINING SEVERAL INDEPENDENT TESTS

Table 2: Normal Distribution $n = 100$, $nrep = 10,000$, $\alpha = 0.05$

μ	KS	P(1)	P(2)	P(3)	P(4)	P(5)	INV-CHI	INV-NORM	LOGIT
4.0	1.0000	0.9947	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.1	1.0000	0.9713	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.2	1.0000	0.8992	0.9969	0.9995	0.9998	1.0000	1.0000	1.0000	1.0000
4.3	1.0000	0.7659	0.9681	0.9936	0.9983	0.9995	1.0000	1.0000	1.0000
4.4	0.9995	0.5823	0.8658	0.9480	0.9782	0.9888	1.0000	0.9999	0.9997
4.5	0.9811	0.4042	0.6596	0.7916	0.8662	0.9098	0.9978	0.9934	0.9876
4.6	0.7915	0.2547	0.4307	0.5320	0.6075	0.6601	0.9298	0.8763	0.8212
4.7	0.4084	0.1572	0.2358	0.2899	0.3263	0.3569	0.6109	0.5204	0.4148
4.8	0.1627	0.0980	0.1179	0.1351	0.1495	0.1519	0.2440	0.2052	0.1353
4.9	0.0834	0.0633	0.0687	0.0705	0.0683	0.0699	0.0795	0.0729	0.0522
5.0	0.0765	0.0475	0.0477	0.0466	0.0508	0.0539	0.0524	0.0511	0.0498
5.1	0.0864	0.0658	0.0668	0.0684	0.0684	0.0717	0.0848	0.0765	0.0545
5.2	0.1587	0.0997	0.1178	0.1314	0.1456	0.1548	0.2423	0.2063	0.1364
5.3	0.4102	0.1619	0.2367	0.2925	0.3360	0.3651	0.6093	0.5189	0.4105
5.4	0.8004	0.2626	0.4307	0.5503	0.6245	0.6747	0.9314	0.8825	0.8274
5.5	0.9805	0.4072	0.6626	0.7906	0.8600	0.8998	0.9975	0.9931	0.9872
5.6	0.9997	0.5881	0.8672	0.9507	0.9768	0.9875	1.0000	1.0000	0.9999
5.7	1.0000	0.7548	0.9661	0.9945	0.9991	0.9996	1.0000	1.0000	1.0000
5.8	1.0000	0.8929	0.9957	0.9996	1.0000	1.0000	1.0000	1.0000	1.0000
5.9	1.0000	0.9674	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6.0	1.0000	0.9940	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figure 1: Normal Distribution $n = 10$, $nrep = 10,000$, $\alpha = 0.05$

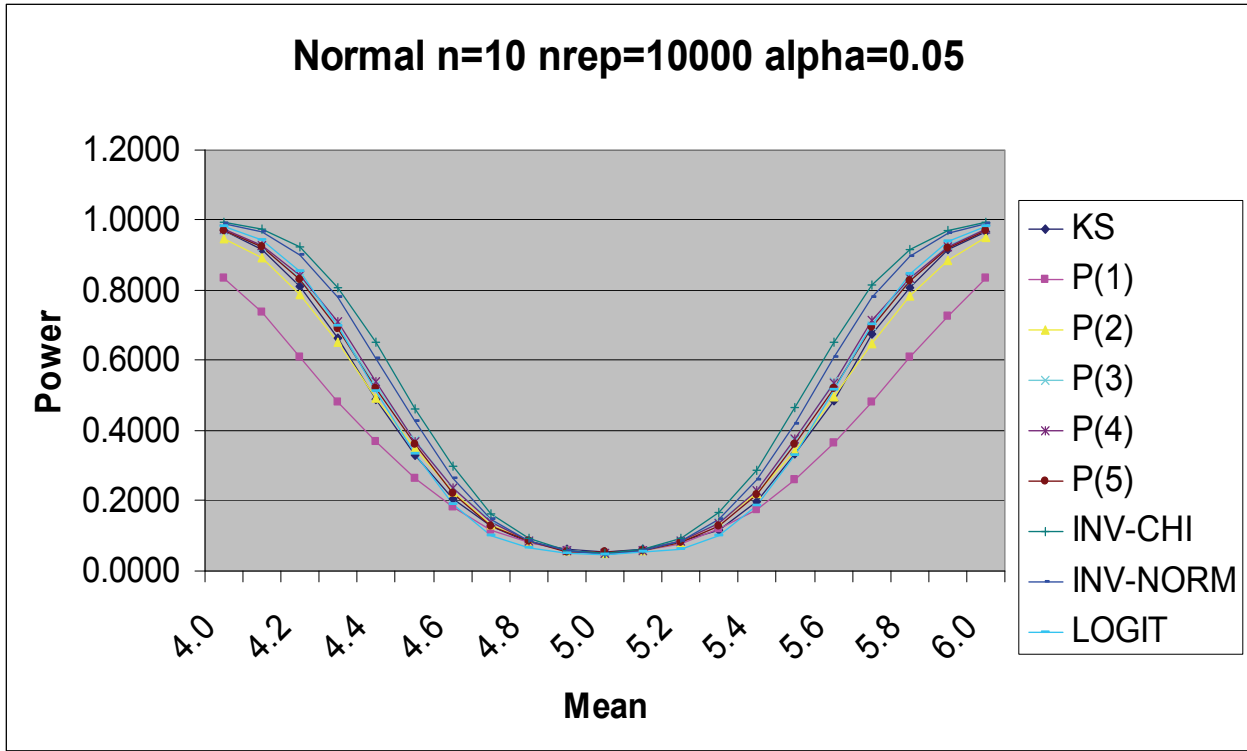
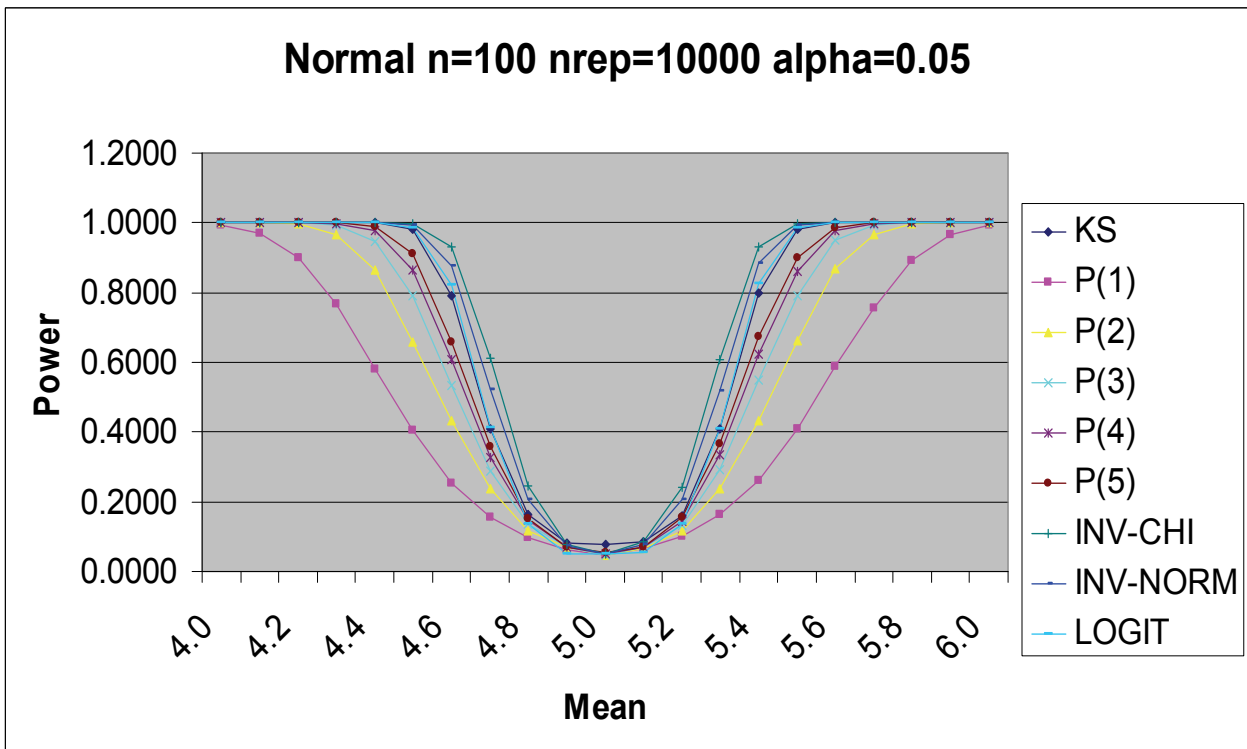


Figure 2: Normal Distribution $n = 100$, $nrep = 10,000$, $\alpha = 0.05$



COMPARISON OF TESTS FOR COMBINING SEVERAL INDEPENDENT TESTS

Table 3: Exponential Distribution $n = 10$, $nrep = 10,000$, $\alpha = 0.05$

μ	KS	P(1)	P(2)	P(3)	P(4)	P(5)	INV-CHI	INV-NORM	LOGIT
4.0	0.8962	0.6269	0.8249	0.8834	0.9039	0.8990	0.9560	0.9503	0.9141
4.1	0.7830	0.5042	0.7091	0.7804	0.7964	0.7944	0.8823	0.8701	0.8064
4.2	0.6375	0.3978	0.5681	0.6391	0.6609	0.6545	0.7619	0.7394	0.6505
4.3	0.4689	0.3033	0.4425	0.4964	0.5089	0.5020	0.6033	0.5731	0.4712
4.4	0.3293	0.2340	0.3229	0.3611	0.3698	0.3545	0.4451	0.4164	0.3164
4.5	0.2167	0.1786	0.2253	0.2452	0.2496	0.2384	0.3017	0.2814	0.1962
4.6	0.1366	0.1222	0.1479	0.1556	0.1548	0.1486	0.1811	0.1712	0.1133
4.7	0.0940	0.0897	0.1039	0.1053	0.1014	0.1006	0.1185	0.1103	0.0742
4.8	0.0716	0.0670	0.0686	0.0722	0.0685	0.0686	0.0743	0.0711	0.0577
4.9	0.0636	0.0518	0.0578	0.0561	0.0558	0.0517	0.0553	0.0553	0.0510
5.0	0.0572	0.0481	0.0524	0.0546	0.0526	0.0526	0.0557	0.0554	0.0511
5.1	0.0599	0.0552	0.0589	0.0564	0.0536	0.0554	0.0562	0.0536	0.0493
5.2	0.0713	0.0651	0.0669	0.0683	0.0774	0.0682	0.0749	0.0695	0.0581
5.3	0.0904	0.0884	0.0953	0.0999	0.0953	0.0926	0.1123	0.1048	0.0746
5.4	0.1275	0.1289	0.1474	0.1496	0.1422	0.1382	0.1748	0.1532	0.1086
5.5	0.1754	0.1734	0.2067	0.2141	0.2069	0.1946	0.2586	0.2286	0.1647
5.6	0.2369	0.2236	0.2801	0.2966	0.2804	0.2650	0.3588	0.3097	0.2339
5.7	0.3293	0.2902	0.3807	0.3971	0.3825	0.3596	0.4907	0.4345	0.3528
5.8	0.4445	0.3738	0.4886	0.5048	0.5007	0.4770	0.6202	0.5643	0.4774
5.9	0.5556	0.4670	0.5907	0.6278	0.6209	0.5966	0.7407	0.6935	0.6100
6.0	0.6666	0.5468	0.6933	0.7277	0.7274	0.6898	0.8314	0.7867	0.7243

Table 4: Exponential Distribution $n = 100$, $nrep = 10,000$, $\alpha = 0.05$

μ	KS	P(1)	P(2)	P(3)	P(4)	P(5)	INV-CHI	INV-NORM	LOGIT
4.0	1.0000	0.8712	0.9963	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
4.1	1.0000	0.7470	0.9732	0.9968	0.9991	0.9998	1.0000	1.0000	1.0000
4.2	1.0000	0.6055	0.9049	0.9717	0.9895	0.9962	1.0000	1.0000	1.0000
4.3	0.9992	0.4563	0.7605	0.8903	0.9437	0.9691	1.0000	0.9999	0.9996
4.4	0.9808	0.3351	0.5863	0.7254	0.8072	0.8605	0.9979	0.9938	0.9866
4.5	0.8378	0.2387	0.3969	0.5073	0.5945	0.6458	0.9436	0.9056	0.8500
4.6	0.5128	0.1590	0.2455	0.3097	0.3630	0.4033	0.7017	0.6235	0.5127
4.7	0.2440	0.1077	0.1451	0.1740	0.1910	0.2106	0.3639	0.3086	0.2167
4.8	0.1166	0.0728	0.0876	0.0952	0.1012	0.1078	0.1483	0.1276	0.0835
4.9	0.0846	0.0565	0.0608	0.0578	0.0613	0.0613	0.0736	0.0725	0.0558
5.0	0.0804	0.0487	0.0495	0.0475	0.0503	0.0527	0.0519	0.0506	0.0485
5.1	0.0796	0.0534	0.0551	0.0619	0.0598	0.0624	0.0670	0.0649	0.0483
5.2	0.1134	0.0820	0.0890	0.0908	0.0971	0.1043	0.1468	0.1276	0.0820
5.3	0.2106	0.1198	0.1539	0.1738	0.1972	0.2111	0.3360	0.2789	0.1877
5.4	0.4033	0.1845	0.2568	0.3201	0.3585	0.3834	0.6240	0.5311	0.4205
5.5	0.6834	0.2675	0.4090	0.5053	0.5665	0.6185	0.8805	0.7972	0.7174
5.6	0.9024	0.3724	0.5874	0.7071	0.7770	0.8231	0.9815	0.9551	0.9275
5.7	0.9852	0.5028	0.7528	0.8647	0.9114	0.9397	0.9986	0.9956	0.9913
5.8	0.9984	0.6325	0.8851	0.9519	0.9769	0.9869	1.0000	0.9998	0.9995
5.9	0.9999	0.7602	0.9566	0.9879	0.9954	0.9985	1.0000	1.0000	1.0000
6.0	1.0000	0.8643	0.9867	0.9981	0.9994	1.0000	1.0000	1.0000	1.0000

COMPARISON OF TESTS FOR COMBINING SEVERAL INDEPENDENT TESTS

Figure 3: Exponential Distribution $n = 10$, $nrep = 10,000$, $\alpha = 0.05$

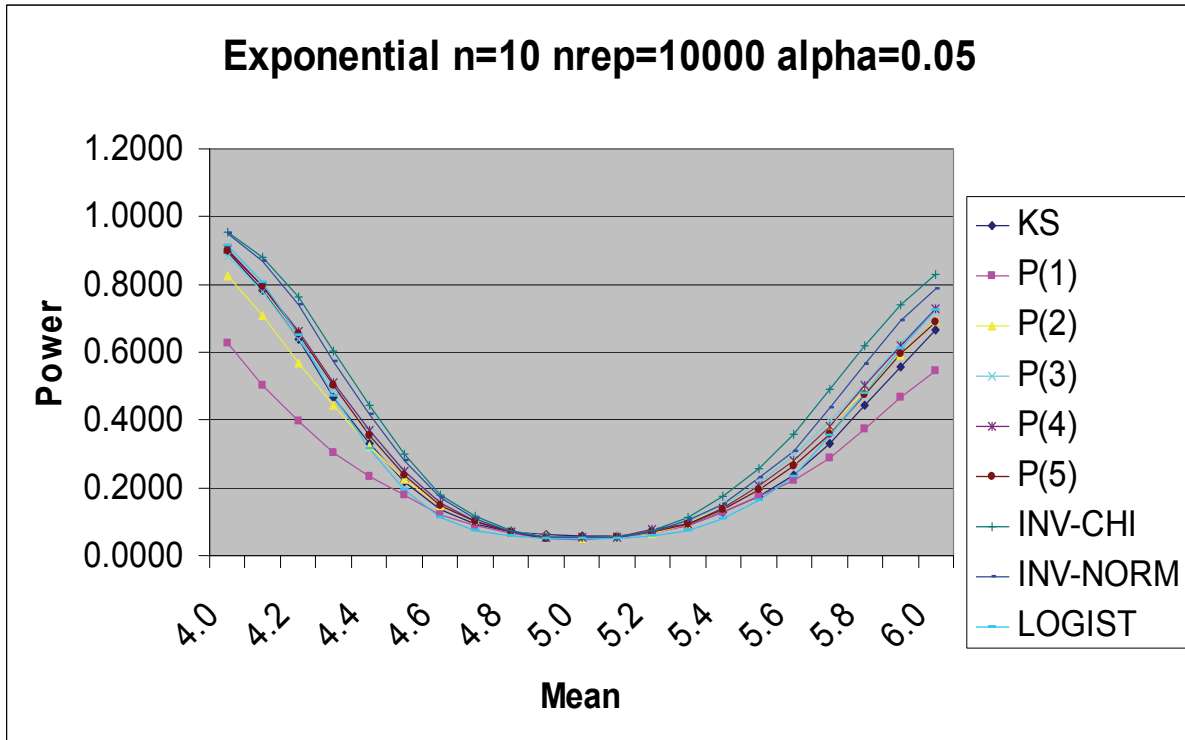


Figure 4: Exponential Distribution $n = 100$, $nrep = 10,000$, $\alpha = 0.05$

