A TEST OF A SAMPLE VARIANCE BASED ON BOTH TAIL ENDS OF THE DISTRIBUTION

By John W. Fertig

With the assistance of Elizabeth A. Proehl[1]

(1) Introduction

In testing the hypothesis, say H_0, that an observed sample E of size N has been drawn from a normal population for which the standard deviation, σ, has a particular value, σ_0, one may form the ratio

$$v = \underset{i=1}{\overset{N}{S}} (x_i - m)^2/\sigma_0^2 = \frac{Nd^2}{\sigma_0^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(I)}$$

if the population mean m be known, or

$$v' = \underset{i=1}{\overset{N}{S}} (x_i - \bar{x})^2/\sigma_0^2 = \frac{Ns^2}{\sigma_0^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(II)}$$

where $\bar{x}$ is the sample mean, if the population mean be unknown. The probability of obtaining a larger (or smaller) value of v or v' than that observed may readily be obtained from the appropriate tail area of the χ^2 distribution with $n = N$ or $n = (N - 1)$ degrees of freedom respectively. The alternative hypotheses to H_0 concerning the normal populations from which the sample may have been drawn assign different values to σ and form a set of hypotheses, Ω. The members of Ω may be classed according to whether they specify $\sigma > \sigma_0$, or $\sigma < \sigma_0$. The practice of regarding only one tail of the distribution, the upper or lower depending on whether $v > N$ or $v < N$, is tantamount to accepting as admissible alternatives to H_0 only one of the classes of Ω.

The alternatives may sometimes be limited to one class or the other through some a priori knowledge, or the problem may be such that only one of the classes is relevant. However, since this is not generally the case, some method of considering all of the alternatives is needed. When testing hypotheses concerning the mean of the sampled population, the problem is quite simple, since the distribution of means is symmetrical. Thus, the "corresponding" value to any positive deviation, $(\bar{x} - m)$, is the negative deviation of the same magnitude. Merely doubling the tail area pertaining to either of the deviations will serve to take account of both classes of alternatives, i.e., those in which $m > m_0$ and those in which $m < m_0$. The problem is more difficult in the case of v or v',

[1] From the Memorial Foundation for Neuro-Endocrine Research and the Research Service of the Worcester State Hospital, Worcester, Massachusetts.

since the distribution is not symmetrical. In addition to the value of v or v' pertaining to the observed sample we require a "corresponding" value at the other end of the distribution. The definition of "corresponding" which is accepted will determine the required value. There may be a number of such definitions but not all of these will be equally acceptable. The value of v which delimits an equal tail area specifies one of the possible definitions of "corresponding." Another definition would require that the ordinates at the two values of v be equal.

The Neyman and Pearson Approach. Generalized procedures for testing statistical hypotheses have been elaborated in recent years by J. Neyman and E. S. Pearson (1–5). These have considerable philosophical appeal and will be traced as a basis of solution of the immediate problem. A test of a hypothesis H_0 consists essentially of a rule for rejecting H_0 when the observed sample E falls within a suitable critical region w of the N-dimensioned sample space W, and of accepting H_0 when E falls in $(W - w)$. In testing any hypothesis two types of error may be made:

i) H_0 may be rejected when it is true;

ii) H_0 may be accepted when some alternative hypothesis, H_i, is true.

Errors of the first kind may be considered "equivalent" since, if a true hypothesis is to be rejected, it is immaterial which one is chosen. Furthermore, the first type of error can be controlled through our choice of the size of w, say α. The size of w represents the probability of a sample E being an element of w when the hypothesis H_0 is true. This probability may be designated briefly as $P\{E \epsilon w \mid H_0\}$. Then

$$P\{E \epsilon w \mid H_0\} = \int \cdots \int_w p(E \mid H_0)\, dx_1 dx_2 \cdots dx_N = \alpha \ldots\ldots \text{(III)}$$

where $p(E \mid H_0)$ is the elementary probability law of the sample when H_0 is true, i.e.,

$$p(E \mid H_0) = p(x_1, x_2, \cdots x_N \mid H_0) \ldots\ldots\ldots\ldots\ldots \text{(IV)}$$

Errors of the second type, however, are not equivalent, since their consequences depend on the difference of the true hypothesis from H_0. The utility of a test of H_0 will depend largely on how it controls the second type of error. Ideally, the selection of a critical region should take into consideration the probabilities á priori of the hypotheses composing Ω. Since these probabilities are generally unknown, tests may be sought which are valid independently of them.

A distinction must be made between simple hypotheses which specify completely the elementary probability law of the sample, $p(E)$, and composite hypotheses which specify the law subject to one or more undetermined parameters.

(2) Simple Hypothesis Concerning Population Variance

A test based on a critical region w_0 may be called independent of the probabilities á priori of the alternative hypotheses if it is more powerful than any other equivalent test for all of the alternative hypotheses (3). An equivalent test is one based on a region w_1 of the same size, α, i.e.,

$$P\{E \epsilon w_0 \mid H_0\} = P\{E \epsilon w_1 \mid H_0\} = \alpha \ldots\ldots\ldots\ldots\ldots \text{(V)}$$

The power of a test based on any critical region, as w_1, is the probability of its rejecting a hypothesis H_0 when some other hypothesis H_i is true. That is, it is the probability of E falling in w_1 when H_i is true. Denote this power by $P\{E \epsilon w_1 \mid H_i\}$. The greater the power of a test, the smaller the risk of the second type of error. If tests as defined above exist, they minimize the probability of the second type of error. Furthermore, the probability of the first type of error is no larger than α. Neyman and Pearson (2) have designated regions satisfying this definition as Best Critical Regions for testing H_0 with regard to the set Ω. If there is no such Best Critical Region, some compromise region must be chosen.

A necessary and sufficient condition for w_0 to be a Best Critical Region with regard to an alternative H_i is that within w_0

$$p(E \mid H_0) \leq kp(E \mid H_i) \ldots\ldots\ldots\ldots\ldots\ldots \text{(VI)}$$

where k is some constant depending on α. If this inequality is true for any H_i, w_0 will be a Best Critical Region for the set Ω.

Neyman and Pearson (2) have shown that in testing the hypothesis that $\sigma = \sigma_0$, when the population mean m is known, there are two Best Critical regions, one pertaining to the class of alternatives for which $\sigma < \sigma_0$ and defined by $v \leq v_1$, the other to the class $\sigma > \sigma_0$ defined by $v \geq v_2$. v_1 and v_2 are values of v so chosen that the size of the critical region shall be α. Although there is no Best Critical Region for all of the alternatives, the choice of a compromise critical region should still depend on its control of the second source of error, that is, on its power for the various alternatives (4). Such a compromise region may be designated as a Good Critical Region. What is needed is a region w_0 of size α defined by the inequalities $v \leq v_1$ and $v \geq v_2$. If v_1 and v_2 are taken as the values cutting off equal tail areas, then the power of the test will be less than α for some values of σ less than σ_0. For those values of σ, H_0 would be accepted more frequently than if it were true. Thus a first requirement for a Good Critical Region is that its power should nowhere be less than α, the value when H_0 is true. Of all such unbiassed Critical Regions of size α, w_0 should then be selected so that its power is everywhere greater than that of any other equivalent unbiassed region.

Critical Regions sufficiently satisfying the above requirements can often be obtained by stipulating that the first derivative of the power function with respect to θ, the parameter under consideration, shall be zero at $\theta = \theta_0$, and that the second shall be a maximum there. Then not only does the probability of the second source of error decrease as we move away from θ_0, but it decreases most rapidly in the vicinity of θ_0. Critical Regions satisfying these conditions are called unbiassed Critical Regions of Type A, (4). Under certain assumptions concerning the nature of the elementary probability law $p(E \mid \theta)$ it can be shown that w_0 is defined by the inequalities $\varphi_1 \leq c_1$ and $\varphi_1 \geq c_2$ where c_1 and c_2 satisfy the conditions

$$\int_{c_1}^{c_2} p(\varphi_1)\, d\varphi_1 = 1 - \alpha \quad \text{(VII)}$$

$$\int_{c_1}^{c_2} \varphi_1 p(\varphi_1)\, d\varphi_1 = 0 \quad \text{(VIII)}$$

where
$$\varphi_1 = \left.\frac{d \log p(E \mid \theta)}{d\theta}\right|_{\theta=\theta_0} \quad \text{(IX)}$$

and $p(\varphi_1)$ is the distribution function of φ_1.

In applying these results to the testing of the hypothesis that $\sigma = \sigma_0^2$ when the population mean is known,

$$\varphi_1 = (v - N)/\sigma_0 \quad \text{(X)}$$

Obviously $p(v)$, the distribution of v, may be considered instead of $p(\varphi_1)$. w_0 is defined by the inequalities $v \leq v_1$ and $v \geq v_2$ where

$$\int_0^{v_1} p(v)\, dv + \int_{v_2}^{\infty} p(v)\, dv = \alpha_1 + \alpha_2 = \alpha \quad \text{(XI)}$$

$$\int_{v_1}^{v_2} (v - N)p(v)\, dv = v^{N/2}e^{-v/2}\Big|_{v_1}^{v_2} = 0 \quad \text{(XII)}$$

w_0 so defined is also of type A_1, that is, its power curve lies everywhere above that of any other equivalent region, vanishing in the first derivative at $\sigma = \sigma_0$, (4).

The use of w_0 as the appropriate critical region is equivalent to the use of r as a test criterion, where

$$v^{N/2}e^{-\frac{1}{2}v} = r) \quad \text{XIII)}$$

That is, a value of v yielding the same r as the observed v may be taken as the corresponding value. Reference to the appropriate tables and summing of the two tail areas gives P_r, the probability of obtaining a smaller value of r when H_0 is true. H_0 may be rejected if P_r is less than some previously fixed number, say α. If the distribution of r could be evaluated the necessity of dealing with two values of v would be obviated.

The criterion r is equivalent to that deduced by the use of maximum likelihood ratios (6). Thus,

$$p(E \mid \sigma^2) = (2\pi\sigma^2)^{-N/2} e^{-\overset{N}{\underset{i=1}{S}}(x_i - m)^2/2\sigma^2} \quad \text{(XIV)}$$

---

[2] The solution is the same in terms of σ^2.

Maximizing $p(E \mid \sigma^2)$ for fixed E and all possible σ^2 we have

$$p_{\text{max.}}(E \mid \sigma^2) = N^{N/2}\left[2\pi \underset{i=1}{\overset{N}{S}} (x_i - m)^2\right]^{-N/2} e^{-N/2} \quad \text{(XV)}$$

$$\lambda = \frac{p(E \mid \sigma_0^2)}{p_{\text{max.}}(E \mid \sigma^2)} = N^{-N/2} v^{N/2} e^{-\frac{1}{2}(v-N)} \quad \text{(XVI)}$$

$$= N^{-N/2} e^{N/2} r \quad \text{(XVII)}$$

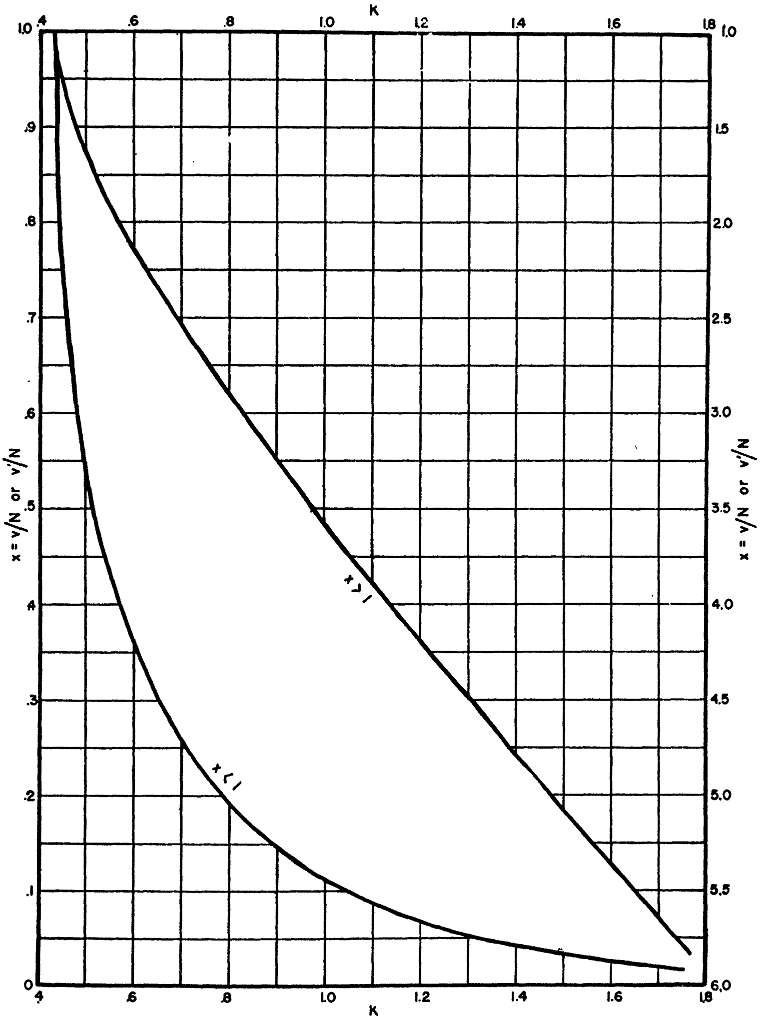

FIG. 1. Graph of Equation $x - \log_e x = k \log_e 10$

The h^{th} moment coefficient of λ about zero, $\mu_h'(\lambda)$, is given by

$$\mu_h'(\lambda) = \frac{\Gamma\left[\frac{N(1+h)}{2}\right]}{\Gamma(N/2)} (2e/N)^{hN/2}(1+h)^{-N(1+h)/2} \quad \text{(XVIII)}$$

TABLE I

Probability that a sample has been drawn from a normal population with a specified variance or standard deviation

Degrees of Freedom, n

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0.435	.9724	.9581	.9473	.9383	.9305	.9235	.9171	.9111	.9055	.9002	.8952	.8905	.8859	.8815	.8773	.8732	.8693	.8655	.8617	.8581	.8546	.8512	.8478	.8445	.8413
0.440	.9217	.8812	.8509	.8259	.8042	.7848	.7671	.7508	.7357	.7215	.7081	.6954	.6833	.6717	.6606	.6500	.6398	.6299	.6204	.6112	.6023	.5936	.5853	.5771	.5692
0.445	.8928	.8377	.7968	.7632	.7341	.7083	.6850	.6636	.6438	.6253	.6080	.5917	.5762	.5616	.5476	.5343	.5215	.5093	.4975	.4862	.4753	.4649	.4547	.4450	.4355
0.450	.8704	.8041	.7552	.7151	.6808	.6505	.6232	.5983	.5754	.5542	.5344	.5159	.4984	.4820	.4664	.4516	.4375	.4241	.4113	.3990	.3873	.3760	.3652	.3549	.3449
0.455	.8513	.7758	.7203	.6752	.6367	.6029	.5726	.5452	.5202	.4971	.4756	.4557	.4370	.4195	.4030	.3874	.3726	.3587	.3454	.3328	.3208	.3094	.2985	.2881	.2781
0.460	.8346	.7510	.6899	.6405	.5987	.5621	.5296	.5003	.4737	.4492	.4267	.4059	.3865	.3683	.3514	.3355	.3205	.3064	.2931	.2806	.2687	.2574	.2467	.2366	.2269
0.465	.8194	.7287	.6628	.6098	.5651	.5263	.4920	.4613	.4335	.4082	.3850	.3636	.3439	.3255	.3084	.2925	.2776	.2637	.2506	.2383	.2267	.2158	.2055	.1958	.1866
0.470	.8055	.7083	.6382	.5821	.5350	.4944	.4587	.4270	.3984	.3725	.3489	.3273	.3074	.2890	.2721	.2563	.2417	.2281	.2154	.2035	.1924	.1819	.1722	.1630	.1544
0.475	.7926	.6896	.6156	.5568	.5077	.4657	.4289	.3963	.3672	.3410	.3172	.2956	.2758	.2576	.2409	.2255	.2113	.1981	.1859	.1745	.1639	.1541	.1449	.1364	.1284
0.480	.7805	.6721	.5946	.5335	.4827	.4395	.4019	.3688	.3393	.3130	.2892	.2677	.2481	.2303	.2140	.1990	.1853	.1726	.1610	.1502	.1402	.1310	.1224	.1145	.1071
0.485	.7692	.6557	.5751	.5119	.4597	.4155	.3773	.3438	.3142	.2879	.2643	.2430	.2238	.2064	.1906	.1761	.1630	.1509	.1398	.1296	.1203	.1117	.1037	.0964	.0897
0.490	.7583	.6402	.5569	.4918	.4384	.3934	.3547	.3211	.2915	.2653	.2420	.2211	.2023	.1854	.1701	.1562	.1436	.1322	.1217	.1122	.1035	.0955	.0881	.0814	.0753
0.495	.7481	.6256	.5397	.4729	.4185	.3729	.3340	.3003	.2708	.2449	.2219	.2015	.1832	.1668	.1521	.1388	.1269	.1160	.1062	.0973	.0892	.0818	.0750	.0689	.0633
0.500	.7382	.6117	.5234	.4552	.4000	.3539	.3148	.2812	.2519	.2263	.2038	.1838	.1661	.1503	.1362	.1236	.1122	.1020	.0928	.0845	.0770	.0702	.0640	.0584	.0534
0.510	.7197	.5857	.4933	.4228	.3663	.3197	.2806	.2473	.2188	.1940	.1725	.1537	.1372	.1226	.1097	.0983	.0882	.0792	.0712	.0640	.0577	.0519	.0468	.0422	.0381
0.520	.7025	.5619	.4660	.3937	.3364	.2897	.2509	.2183	.1907	.1670	.1466	.1290	.1138	.1004	.0888	.0786	.0697	.0618	.0549	.0488	.0434	.0386	.0344	.0307	.0273
0.530	.6864	.5399	.4411	.3674	.3097	.2632	.2250	.1933	.1667	.1442	.1251	.1087	.0947	.0826	.0721	.0631	.0553	.0484	.0425	.0373	.0328	.0289	.0254	.0224	.0197
0.540	.6713	.5194	.4181	.3435	.2856	.2396	.2023	.1716	.1461	.1248	.1070	.0918	.0790	.0681	.0588	.0508	.0439	.0381	.0330	.0286	.0249	.0216	.0188	.0164	.0143
0.550	.6570	.5002	.3969	.3216	.2639	.2186	.1822	.1526	.1284	.1083	.0917	.0778	.0661	.0563	.0480	.0410	.0351	.0300	.0257	.0221	.0189	.0163	.0140	.0120	.0103
0.560	.6434	.4822	.3772	.3015	.2442	.1997	.1643	.1360	.1130	.0942	.0787	.0660	.0554	.0466	.0393	.0332	.0280	.0237	.0201	.0170	.0144	.0123	.0104	.0089	.0075
0.570	.6304	.4652	.3588	.2830	.2263	.1827	.1485	.1213	.0996	.0820	.0678	.0561	.0466	.0387	.0322	.0269	.0224	.0188	.0157	.0132	.0110	.0093	.0078	.0065	.0055
0.580	.6180	.4492	.3417	.2659	.2099	.1673	.1343	.1084	.0879	.0715	.0584	.0478	.0392	.0322	.0265	.0218	.0180	.0149	.0123	.0102	.0084	.0070	.0058	.0048	.0040
0.590	.6061	.4340	.3256	.2501	.1949	.1534	.1217	.0970	.0777	.0625	.0504	.0407	.0330	.0268	.0218	.0177	.0145	.0118	.0097	.0079	.0065	.0053	.0044	.0036	.0029
0.600	.5946	.4195	.3105	.2354	.1811	.1408	.1103	.0869	.0688	.0546	.0435	.0348	.0278	.0223	.0180	.0145	.0117	.0094	.0076	.0061	.0050	.0040	.0033	.0027	.0022
0.610	.5836	.4057	.2963	.2217	.1685	.1294	.1001	.0779	.0609	.0478	.0376	.0297	.0235	.0186	.0148	.0118	.0094	.0075	.0060	.0048	.0038	.0031	.0025	.0020	.0016
0.620	.5730	.3926	.2829	.2090	.1568	.1190	.0909	.0699	.0540	.0419	.0326	.0254	.0199	.0156	.0122	.0096	.0076	.0060	.0047	.0037	.0029	.0023	.0018	.0015	.0012
0.630	.5627	.3801	.2702	.1971	.1461	.1094	.0826	.0628	.0479	.0367	.0282	.0218	.0168	.0130	.0101	.0079	.0061	.0048	.0037	.0029	.0023	.0018	.0014	.0011	.0009
0.640	.5528	.3682	.2583	.1860	.1362	.1008	.0752	.0564	.0426	.0322	.0245	.0187	.0143	.0109	.0084	.0064	.0049	.0038	.0029	.0023	.0018	.0014	.0010	.0008	.0006

0.650	.5432	.3567	.2470	.1757	.1270	.0928	.0684	.0508	.0378	.0283	.0213	.0160	.0121	.0091	.0069	.0053	.0040	.0030	.0023	.0018	.0014	.0010	.0008	.0006	.0005
0.660	.5339	.3457	.2363	.1659	.1185	.0856	.0623	.0457	.0336	.0249	.0185	.0137	.0103	.0077	.0057	.0043	.0032	.0024	.0018	.0014	.0010	.0008	.0006	.0005	.0003
0.670	.5249	.3352	.2261	.1568	.1106	.0789	.0568	.0411	.0299	.0219	.0160	.0118	.0087	.0064	.0048	.0035	.0026	.0020	.0015	.0011	.0008	.0006	.0005	.0003	.0003
0.680	.5161	.3251	.2165	.1483	.1033	.0728	.0518	.0371	.0267	.0193	.0140	.0101	.0074	.0054	.0040	.0029	.0021	.0016	.0012	.0008	.0006	.0005	.0003	.0003	.0002
0.690	.5076	.3154	.2073	.1403	.0965	.0672	.0472	.0334	.0237	.0169	.0121	.0087	.0063	.0045	.0033	.0024	.0017	.0013	.0009	.0007	.0005	.0004	.0003	.0002	.0001
0.700	.4993	.3060	.1986	.1327	.0902	.0621	.0431	.0301	.0212	.0149	.0106	.0075	.0053	.0038	.0027	.0020	.0014	.0010	.0007	.0005	.0004	.0003	.0003	.0002	.0001
0.750	.4609	.2642	.1610	.1011	.0647	.0419	.0274	.0181	.0120	.0080	.0053	.0036	.0024	.0016	.0011	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000	.0000	.0000
0.800	.4268	.2292	.1312	.0776	.0468	.0286	.0176	.0109	.0068	.0043	.0027	.0017	.0011	.0007	.0004	.0003	.0002	.0001	.0001	.0000	.0000	.0000			
0.850	.3962	.1995	.1074	.0598	.0339	.0195	.0114	.0067	.0039	.0023	.0014	.0008	.0005	.0003	.0002	.0001	.0001	.0000	.0000						
0.900	.3686	.1742	.0883	.0463	.0248	.0134	.0074	.0041	.0023	.0013	.0007	.0004	.0002	.0001	.0001	.0000	.0000								
0.950	.3435	.1525	.0727	.0359	.0181	.0093	.0048	.0025	.0013	.0007	.0004	.0002	.0001	.0001	.0000										
1.000	.3205	.1338	.0601	.0280	.0133	.0064	.0031	.0015	.0008	.0004	.0002	.0001	.0000	.0000											
1.050	.2994	.1175	.0497	.0218	.0098	.0045	.0020	.0010	.0004	.0002	.0001	.0000													
1.100	.2800	.1034	.0412	.0170	.0072	.0031	.0013	.0006	.0003	.0001	.0001														
1.150	.2621	.0911	.0342	.0133	.0053	.0022	.0009	.0004	.0002	.0001	.0000														
1.200	.2455	.0803	.0284	.0105	.0039	.0015	.0006	.0002	.0001	.0000															
1.250	.2301	.0709	.0236	.0082	.0029	.0011	.0004	.0001	.0001																
1.300	.2158	.0626	.0197	.0064	.0022	.0007	.0003	.0001	.0000																
1.350	.2024	.0553	.0164	.0051	.0016	.0005	.0002	.0001																	
1.400	.1900	.0490	.0137	.0040	.0012	.0004	.0001	.0000																	
1.450	.1785	.0433	.0114	.0031	.0009	.0003	.0001																		
1.500	.1677	.0384	.0096	.0025	.0007	.0002	.0000																		
1.800	.1159	.0187	.0033	.0006	.0001	.0000																			
2.100	.0807	.0092	.0011	.0001	.0000																				
2.400	.0564	.0045	.0004	.0000																					
2.700	.0395	.0022	.0001																						
3.000	.0278	.0011	.0000																						

TABLE I—*Concluded*

k	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
0.435	.8382	.8351	.8321	.8291	.8262	.8234	.8206	.8178	.8151	.8124	.8098	.8072	.8046	.8021	.7996	.7972	.7947	.7923	.7900	.7876	.7853	.7831	.7808	.7786	.7764
0.440	.5614	.5539	.5466	.5394	.5324	.5256	.5189	.5124	.5060	.4998	.4936	.4876	.4818	.4760	.4704	.4648	.4594	.4540	.4488	.4437	.4386	.4336	.4287	.4239	.4192
0.445	.4263	.4175	.4089	.4005	.3924	.3846	.3769	.3695	.3623	.3553	.3484	.3417	.3352	.3289	.3227	.3167	.3108	.3051	.2995	.2940	.2887	.2835	.2784	.2734	.2685
0.450	.3353	.3261	.3172	.3086	.3003	.2923	.2845	.2771	.2698	.2628	.2561	.2495	.2432	.2370	.2310	.2252	.2196	.2142	.2089	.2037	.1988	.1939	.1892	.1846	.1802
0.455	.2686	.2595	.2507	.2424	.2343	.2266	.2192	.2120	.2052	.1986	.1922	.1861	.1802	.1745	.1691	.1638	.1587	.1538	.1490	.1445	.1401	.1358	.1317	.1277	.1238
0.460	.2177	.2090	.2006	.1927	.1851	.1779	.1710	.1644	.1580	.1520	.1462	.1407	.1354	.1303	.1254	.1208	.1163	.1120	.1079	.1039	.1002	.0965	.0930	.0896	.0864
0.465	.1779	.1697	.1619	.1545	.1475	.1409	.1346	.1286	.1229	.1174	.1123	.1074	.1027	.0982	.0940	.0899	.0861	.0824	.0789	.0755	.0723	.0693	.0664	.0636	.0609
0.470	.1463	.1387	.1315	.1247	.1183	.1123	.1066	.1012	.0962	.0914	.0868	.0825	.0784	.0746	.0709	.0675	.0642	.0611	.0581	.0553	.0527	.0501	.0477	.0455	.0433
0.475	.1209	.1139	.1073	.1012	.0954	.0900	.0849	.0801	.0757	.0714	.0675	.0638	.0602	.0569	.0538	.0509	.0481	.0455	.0431	.0407	.0386	.0365	.0345	.0327	.0310
0.480	.1002	.0939	.0879	.0824	.0772	.0724	.0679	.0637	.0598	.0561	.0527	.0495	.0465	.0437	.0410	.0386	.0362	.0341	.0320	.0301	.0284	.0267	.0251	.0236	.0222
0.485	.0834	.0776	.0723	.0673	.0627	.0584	.0545	.0508	.0474	.0442	.0413	.0385	.0360	.0336	.0314	.0293	.0274	.0256	.0239	.0224	.0209	.0196	.0183	.0171	.0160
0.490	.0696	.0644	.0596	.0551	.0511	.0473	.0439	.0406	.0377	.0350	.0324	.0301	.0279	.0259	.0241	.0224	.0208	.0193	.0179	.0167	.0155	.0144	.0134	.0125	.0116
0.495	.0582	.0535	.0492	.0453	.0417	.0384	.0354	.0326	.0300	.0277	.0256	.0236	.0218	.0201	.0185	.0171	.0158	.0146	.0135	.0125	.0115	.0106	.0098	.0091	.0084
0.500	.0487	.0446	.0407	.0373	.0341	.0312	.0286	.0262	.0240	.0220	.0202	.0185	.0170	.0156	.0143	.0131	.0120	.0111	.0102	.0093	.0086	.0079	.0072	.0067	.0061
0.510	.0344	.0311	.0281	.0254	.0229	.0208	.0188	.0170	.0154	.0140	.0126	.0115	.0104	.0094	.0085	.0078	.0070	.0064	.0058	.0053	.0048	.0043	.0039	.0036	.0033
0.520	.0244	.0218	.0194	.0174	.0155	.0139	.0124	.0111	.0099	.0089	.0080	.0071	.0064	.0057	.0051	.0046	.0041	.0037	.0033	.0030	.0027	.0024	.0022	.0019	.0017
0.530	.0174	.0153	.0135	.0119	.0106	.0093	.0082	.0073	.0064	.0057	.0051	.0045	.0040	.0035	.0031	.0028	.0024	.0022	.0019	.0017	.0015	.0013	.0012	.0011	.0009
0.540	.0124	.0108	.0095	.0082	.0072	.0063	.0055	.0048	.0042	.0037	.0032	.0028	.0025	.0022	.0019	.0017	.0014	.0013	.0011	.0010	.0009	.0008	.0007	.0006	.0005
0.550	.0089	.0077	.0066	.0057	.0049	.0043	.0037	.0032	.0027	.0024	.0021	.0018	.0015	.0013	.0011	.0010	.0009	.0007	.0006	.0006	.0005	.0004	.0004	.0003	.0003
0.560	.0064	.0055	.0047	.0040	.0034	.0029	.0025	.0021	.0018	.0015	.0013	.0011	.0010	.0008	.0007	.0006	.0005	.0004	.0004	.0003	.0003	.0002	.0002	.0002	.0001
0.570	.0046	.0039	.0033	.0028	.0023	.0020	.0017	.0014	.0012	.0010	.0008	.0007	.0006	.0005	.0004	.0004	.0003	.0003	.0002	.0002	.0002	.0001	.0001	.0001	.0001
0.580	.0033	.0028	.0023	.0019	.0016	.0013	.0011	.0009	.0008	.0006	.0005	.0005	.0004	.0003	.0003	.0002	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0000
0.590	.0024	.0020	.0016	.0013	.0011	.0009	.0008	.0006	.0005	.0004	.0003	.0003	.0002	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000	
0.600	.0017	.0014	.0012	.0009	.0008	.0006	.0005	.0004	.0003	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000				
0.610	.0013	.0010	.0008	.0007	.0005	.0004	.0003	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000							
0.620	.0009	.0007	.0006	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0000	.0000										
0.630	.0007	.0005	.0004	.0003	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000												
0.640	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0000														
0.650	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000															
0.660	.0003	.0002	.0001	.0001	.0001	.0001	.0000	.0000																	
0.670	.0002	.0001	.0001	.0001	.0001	.0000																			
0.680	.0001	.0001	.0001	.0001	.0000																				
0.690	.0001	.0001	.0001	.0000																					
0.700	.0001	.0001	.0000																						
0.750	.0000	.0000																							

For N infinite, $(-2\log_e\lambda)$ will be distributed as χ^2 with one degree of freedom. For finite values of N, however, we have not been able to evaluate the distribution of λ, although the distribution of the Incomplete Beta Function serves as a good approximation. Approximate distributions for several values of N have been obtained. P_λ, the probability of obtaining a smaller value of λ than that observed, as obtained from these distributions agrees well with the sum of the tail areas pertaining to v_1 and v_2 yielding the same value of λ (or r). The construction of tables is simplified by taking (1)

$$\log_{10} \lambda = N/2(\log_{10} e - k) \quad \text{(XIX)}$$

That is,

$$x - \log_e x = k \log_e 10 \quad \text{(XX)}$$

where $x = v/N$. Equation (XX) is independent of N and may be solved once and for all for x, given k.[3] In Figure 1 is plotted the graph of equation (XX). For convenience, the branch of the curve giving the roots greater than unity has been folded back with altered scale from the minimum value of k, $\log_{10}e$, occurring at $x = 1$. Table I was then constructed by multiplying the two values of x for a given k by $(N/2)^{\frac{1}{2}}$, referring to the Tables of the Incomplete Gamma Function (7) with $p = (N - 2)/2$, and adding the resulting two tail areas. The values for the odd numbers above 12 were obtained by interpolating between the even numbers. For $N = 1$, $(x)^{\frac{1}{2}}$ was used as a normal deviate. The values in Table I should be correct to four decimals. Table I is entered with the number of degrees of freedom, n, on which x is based. In the case of the simple hypothesis this is N.

The following may serve as an illustration: Blood urea nitrogen determinations (mg./100 cc.) were made on a sample of 25 schizophrenic patients. The mean was found to be 15.56, the variance, 10.486. Previous investigation of blood urea nitrogen on a large sample of normal control subjects gave a mean of 16.03 and a variance of 20.268, which for the purpose of the example may be considered as the population parameters. Then we may wish to test the hypothesis that the variance of the sampled population, σ^2, is $\sigma_0^2 = 20.268$, knowing the mean of the sampled population to be 16.03. Calculate

$$x = \frac{s^2 + (\bar{x} - m)^2}{\sigma_0^2} = .528$$

Referring to Fig. 1, the value of k is about .505. Turning to Table I with $k = .505$,[4] $n = 25$, P is found to be .0457. We should thus be inclined to reject the hypothesis.

For N small, the area of the tail of the distribution near zero is considerably larger than that at the upper end. As N increases the distribution of v becomes

[3] If the solution were explicit the distribution of λ could easily be deduced from that of x.

[4] k obtained directly from (XX) is .507, corresponding to $P = .0427$.

more and more symmetrical and the two areas approach equality. Even for $N = 50$, however, they are rather unequal, so that merely doubling the area pertaining to the observed v does not give a sufficiently accurate approximation. For $N > 50$ an approximation correct within several units in the third decimal place may be obtained by taking $\sqrt{2N}(\sqrt{x} - 1)$ as a normal deviate. This assumes that the standard deviation is normally distributed with variance $\sigma_0^2/2N$.

(3) Composite Hypothesis Concerning Population Variance

Here H_0 specifies only the value of the parameter $\theta = \theta_0$, leaving undetermined the value of a second parameter, ν. Thus, H_0 consists of a subset, ω, of simple hypotheses, each of which specifies a different value for ν. Any simple hypothesis specifying different values of both parameters, θ and ν, is an alternative to H_0. These alternatives form the set Ω. The elementary probability law determined by H_0 is $p(E \mid H_0) = p(E \mid \theta_0\nu)$, while that determined by an alternative hypothesis H_i is $p(E \mid H_i) = p(E \mid \theta_i\nu_i)$. In testing composite hypotheses the first requirement is to find regions "similar" to W with regard to ν, i.e., such that the chance of rejection of a true hypothesis, $P\{E \epsilon w \mid H_0\}$, equals α for all the values of ν specified by the simple hypotheses composing H_0. A test based on a similar region w_0 may be called independent of the probabilities á priori, if its power with respect to all the alternatives of Ω is greater than that of any other similar region w_1 of the same size, α, (3). Let

$$\varphi_2 = \partial \log p(E \mid \theta\nu)/\partial\nu \big|_{\theta=\theta_0} \qquad \text{(XXI)}$$

Then the equations $\varphi_2 =$ constant will describe hypersurfaces in N-dimensioned space, on one of which the observed E must fall. Under certain assumptions pertaining to the law of elementary probability it can be shown (2) that a necessary and sufficient condition for w to be a similar region is that

$$P\{E \epsilon w(\varphi_2) \mid H_0\} = \alpha P\{E \epsilon W(\varphi_2) \mid H_0\} \qquad \text{(XXII)}$$

for all values of φ_2, where $w(\varphi_2)$ and $W(\varphi_2)$ are parts of the surface $\varphi_2 =$ constant common to w and W respectively. A similar region is then built up of these parts $w(\varphi_2)$ obtaining for the various values of φ_2. The Best Critical Region, w_0, for a particular simple alternative, H_i, must then be composed of pieces, $w_0(\varphi_2)$, maximizing $P\{E \epsilon w_0(\varphi_2) \mid H_i\}$. The problem is the same as for simple hypotheses except that we shall be working in a space $W(\varphi_2)$ of $(N - 1)$ dimensions. $w_0(\varphi_2)$ is defined by the inequality

$$p(E \mid H_i) \geq k(\varphi_2)\, p(E \mid H_0) \qquad \text{(XXIII)}$$

where $k(\varphi_2)$ is some constant depending on α. If $w_0(\varphi_2)$ is the same for all H_i, then w_0 is the Best Critical Region for testing H_0 with respect to Ω.

Neyman and Pearson showed (2) that in testing the composite hypothesis that $\sigma = \sigma_0$ when the population mean is unknown there are two Best Critical Regions corresponding to the class of alternatives $\sigma < \sigma_0$ and $\sigma > \sigma_0$, defined respectively by the inequalities $v' \leq v_1'$ and $v' \geq v_2'$. If the whole set of alternatives, Ω, is to be considered some compromise region must be sought. Dealing with the case where similar regions exist Neyman (5) defines a Critical Region as unbiassed and of Type B if the first derivative of the power function, $P(E \epsilon w \mid H_i)$, with respect to θ vanishes at $\theta = \theta_0$, and if the second derivative at that point is a maximum. Let

$$\varphi_1 = \left. \frac{\partial \log p(E \mid \theta\nu)}{\partial\theta} \right|_{\theta=\theta_0} \quad \text{(XXIV)}$$

Then it can be shown that the desired region will be defined by the inequalities $\varphi_1 \leq k_1(\varphi_2)$ and $\varphi_1 \geq k_2(\varphi_2)$ where $k_1(\varphi_2)$ and $k_2(\varphi_2)$ are determined to satisfy

$$\int_{k_1(\varphi_2)}^{k_2(\varphi_2)} p(\varphi_1\varphi_2)\, d\varphi_1 = (1 - \alpha)p(\varphi_2) \quad \text{(XXV)}$$

and

$$\int_{k_1(\varphi_2)}^{k_2(\varphi_2)} \varphi_1 p(\varphi_1\varphi_2)\, d\varphi_1 = (1 - \alpha) \int_{-\infty}^{\infty} \varphi_1 p(\varphi_1\varphi_2)\, d\varphi_1 \quad \text{(XXVI)}$$

where $p(\varphi_2)$ is the distribution function of φ_2, and $p(\varphi_1\varphi_2)$ is the simultaneous distribution of φ_1 and φ_2.

Applying equations (XXV) and (XXVI) it follows that the appropriate Critical Region is defined by the inequalities $v' \leq v_1'$ and $v' \geq v_2'$ where

$$\alpha = \alpha_1 + \alpha_2 = \int_0^{v_1'} p(v')\, dv' + \int_{v_2'}^{\infty} p(v')\, dv' \quad \text{(XXVII)}$$

and

$$\left. v'^{(N-1)/2} e^{-\frac{1}{2}v'} \right|_{v_1'}^{v_2'} = 0 \quad \text{(XXVIII)}$$

where $p(v')$ is the distribution function of v'.

The use of the unbiassed Critical Region of Type B corresponds to adopting as a criterion

$$v'^{(N-1)/2} e^{-\frac{1}{2}v'} = r' \quad \text{(XXIX)}$$

Since v' derived from a sample of size N is distributed as v derived from a sample of size $(N - 1)$, it follows that r' is equivalent to the r of equation (XIII) based on a sample of size $(N - 1)$. Therefore Table I may also be used for testing the hypothesis that $\sigma = \sigma_0$ whatever be the population mean, by entering with the number of degrees of freedom, $N - 1$.

In the example previously used, compute

$$x = \frac{s^2}{\sigma_0^2} = 0.517$$

From Figure 1, k is approximately .51, corresponding to $P = .0422$.

r' is not the same as the maximum likelihood ratio λ' (6).

$$\lambda' = \frac{p_{\max}(E \mid \sigma_0^2 m)}{p_{\max}(E \mid \sigma^2 m)} = N^{-N/2} v'^{N/2} e^{-\frac{1}{2}(v'-N)} = N^{-N/2} e^{N/2} v'^{\frac{1}{2}} r' \ldots \text{(XXX)}$$

As N becomes infinite the distribution of λ' is the same as that of the λ of (XVI). For $N = 49$, the probabilities corresponding to λ' agree with those using r' to within a unit in the third decimal.

The λ' test is biassed as may be seen in Figure 2 where we have plotted the power of the test based on the region w defined by $v_1' = 3.187$, $v_2' = 22.912$ for which $\alpha = .0436 + .0064 = .0500$, on the assumption that $\sigma_0^2 = 1.0$, for $N = 10$. Although the criterion is biassed it is slightly more sensitive to alternatives

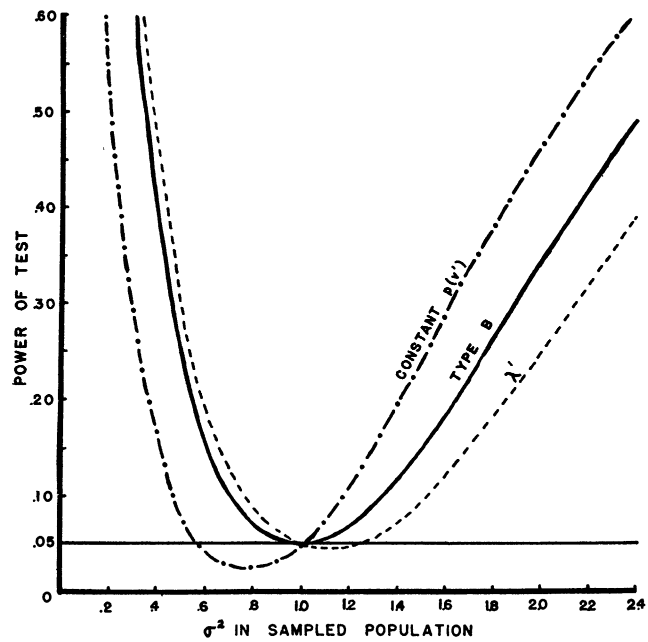

FIG. 2. Comparison of Critical Regions for v'. H_0 Specifies $\sigma_0^2 = 1.0$. $N = 10$.

specifying $\sigma^2 < \sigma_0^2$ than is the unbiassed Critical Region of Type B defined by $v_1' = 2.953$, $v_2' = 20.305$, $\alpha = .0339 + .0161 = .0500$. The criterion of constant distribution, $p(v')$,

$$v'^{(N-3)/2} e^{-\frac{1}{2}v'} = c' \ldots\ldots\ldots\ldots\ldots\ldots \text{(XXXI)}$$

has also been considered. In this case $v_1' = 1.903$, $v_2' = 17.391$, $\alpha = .0071 + .0429 = .0500$. This criterion is biassed for some alternatives specifying $\sigma^2 < \sigma_0^2$, but its power curve lies above that of the unbiassed region for $\sigma^2 > \sigma_0^2$.

Apparently the bias may be shifted at will by changing the exponent of v'. This may be desirable if greater weight is to be given to one class of alternatives. In fact decreasing the exponent of v' to 0 produces the Best Critical Region for the class of alternatives specifying $\sigma^2 > \sigma_0^2$, and defined by $v_1 = 0$, $v_2 = 16.919$ for $\alpha = .0500$. No region can be found giving greater power. On the other hand this region is insensitive to alternatives of the other class. Increasing the exponent indefinitely produces the Best Critical Region for the other class defined by $v_2' = \infty$ and $v_1' = 3.325$ for $\alpha = .0500$.

REFERENCES

(1) Neyman, J. and Pearson, E. S. Biometrika, vol. 20, 1928, pp. 175–241.
(2) Ibid. Phil. Trans. Roy. Soc. A, vol. 231, 1933, pp. 289–337.
(3) Ibid. Proc. Camb. Phil. Soc., vol. 29, 1933, pp. 492–510.
(4) Ibid. Stat. Research Memoirs, vol. 1, 1936, pp. 1–37.
(5) Neyman, J. Bull. d. l. Soc. Math. d. France, vol. 63, 1935, pp. 246–266.
(6) Fertig, J. Biom. Bull., vol. 1, 1936, pp. 45–82.
(7) Tables of the Incomplete Γ-Function, 1934.