The Folded Normal Distribution*

F. C. LEONE[1], L. S. NELSON[2], R. B. NOTTINGHAM[3]

Measurements are frequently recorded without their algebraic sign. As a consequence, the underlying distribution of measurements is replaced by a distribution of absolute measurements. When the underlying distribution is normal, the resulting distribution is called the "folded normal distribution". The authors describe methods for estimating the mean and standard deviation of the normal distribution based on estimates of the mean and standard deviation determined from the folded normal. Tables are provided to assist in the estimation procedure and an example included.

1. INTRODUCTION

In the past few years inspection by variables has found increasing use through efforts to increase inspection efficiency. The usual assumption of normality is valid in many applications, though not in all. One apparently fairly common type of non-normality has recently come to the attention of the authors. Whenever a difference or deviation is measured and the algebraic sign is unknown, disregarded, lost, or otherwise absent, the resulting distribution of these absolute measurements can range in shape from Daniel's "half normal" [l] to the normal distribution as the limit.

Typical of the examples which arise in industrial practice are:

(1) measurement of flatness in which the object is placed convex side up (irrespective of which side this is) and a feeler gauge is used to find the difference between the surface plate and the object.
(2) measurement of the camber (straightness) of mininature radio tube leads (see Sec. 4).
(3) determination of the centrality of the sprocket holes in motion picture film by finding the difference between the distance (Left edge to Left hole) and (Right edge to Right hole). The edge and hole designations are arbitrary because the (unused) film strip sample is not oriented. The distance between the holes is constant since a single die punches the holes simultaneously.

The characteristic which these examples share is that a measurement is recorded without its algebraic sign. The effect of dropping the sign is to add the otherwise negative values to the positive values. Geometrically this amounts to folding the negative side of the distribution onto the positive side. When the underlying distribution (of the algebraic values) is normal, the distribution of the absolute measurements is here described as the ***folded*** normal ***distribution.***

1 Case Institute of Technology, Cleveland, Ohio
2 General Electric Lamp Division, Cleveland, Ohio
North Electric Company, Ohio

* This research was supported in part by the United States Air Force through the Air Force Office of Scientific Research and Development Command.

Daniel's half normal distribution is folded at the mean. In the general case treated here, the "fold" can occur at any distance from the original mean. We propose to consider first the folded normal probability density function (pdf), especially as it relates to the original normal pdf from which it came. We then consider some maximum likelihood estimates, followed by another estimating procedure which is simpler to handle. A table is presented for the estimation of μ and σ (the mean and standard deviation of the original normal) given an estimate of μ_f and σ_f (of the folded normal). A second table gives the areas of the folded normal for values of t_f from 0 to 7 for eighteen different values of μ_f/σ_f.

Finally, an example of real camber data is presented with an estimation of the theoretical distributions.

2. The Folded Normal Distribution

To derive a folded normal probability density function, first consider a normal probability density function

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2}, \qquad -\infty < y < \infty, \tag{1}$$

where μ and σ^2 are the mean and the variance, respectively. Consider now the situation where $x = |y|$, with $x \geq 0$. The probability function has been "folded" at $y = 0$. Generally, in situations such as those previously mentioned, the values of μ and σ are not known. Thus the distance between the mean μ and the fold is unknown.

The probability density function of the resulting folded normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} [e^{-(x-\mu)^2/2\sigma^2} + e^{-(x+\mu)^2/2\sigma^2}], \qquad x \geq 0. \tag{2}$$

By direct integration we obtain as the mean value of x

$$\mu_f = \sqrt{2/\pi}\sigma e^{-\mu^2/2\sigma^2} + \mu[1 - 2F(-\mu/\sigma)], \tag{3}$$

where

$$F(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-t^2/2}\, dt.$$

Similarily σ_f^2, the variance of the folded normal, is given by

$$\sigma_f^2 = \mu^2 + \sigma^2 - \{(2/\pi)^{\frac{1}{2}}\sigma e^{-\mu^2/2\sigma^2} + \mu[1 - 2F(-\mu/\sigma)]\}^2. \tag{4}$$

Maximum likelihood estimates of μ and σ^2 are obtained in the usual manner. The estimating equations which result are

$$\sum_{i=1}^{n} \frac{x_i}{1 + e^{-2\hat{\mu}x_i/\hat{\sigma}^2}} = \left(n\hat{\mu} + \sum_{i=1}^{n} x_i\right)/2 \tag{5}$$

and

$$\sum_{i=1}^{n} \frac{x_i}{1 + e^{-2\hat{\mu}x_i/\hat{\sigma}^2}} = \left[\sum_{i=1}^{n} (x_i + \hat{\mu})^2 - \hat{\sigma}^2 n\right]/4\hat{\mu}. \tag{6}$$

Equating these gives

$$\hat{\sigma}^2 = (1/n) \sum_{i=1}^{n} x_i^2 - \hat{\mu}^2.$$

Maximum likelihood estimation is not practical here because of the complexity of the calculation.

Another approach for estimating μ and σ^2 is to obtain estimates of μ_f and σ_f^2. It is reasonable to use as estimates of μ_f and σ_f^2 the unbiased statistics

$$\bar{x} = (1/n) \sum_{i=1}^{n} x_i \quad \text{and} \quad s^2 = [1/(n-1)] \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

respectively. With these estimates and equations (3) and (4) one can then solve for estimates of μ and σ^2.

From equations (3) and (4) we readily derive

$$\frac{(\mu_f/\sigma_f)^2}{1 + (\mu_f/\sigma_f)^2} = G(\theta) \tag{7}$$

TABLE 1

Values of σ_f and μ Corresponding to Various Values of μ_f/σ_f

	In σ Units			In σ Units	
μ_f/σ_f	σ_f	μ	μ_f/σ_f	σ_f	μ
1.3236	0.6028	0	1.62	0.8758	1.3341
1.325	0.6233	0.2656	1.64	0.8825	1.3689
1.33	0.6468	0.3979	1.66	0.8889	1.4027
1.34	0.6733	0.5170	1.68	0.8949	1.4357
1.35	0.6922	0.5935	1.70	0.9005	1.4678
1.36	0.7076	0.6532	1.75	0.9132	1.5453
1.37	0.7208	0.7034	1.80	0.9242	1.6192
1.38	0.7326	0.7476	1.85	0.9338	1.6900
1.39	0.7433	0.7874	1.90	0.9421	1.7583
1.40	0.7531	0.8238	1.95	0.9494	1.8245
1.41	0.7621	0.8577	2.00	0.9558	1.8888
1.42	0.7706	0.8894	2.05	0.9614	1.9515
1.43	0.7785	0.9194	2.10	0.9663	2.1028
1.44	0.7860	0.9480	2.15	0.9706	2.0728
1.45	0.7930	0.9752	2.20	0.9743	2.1317
1.46	0.7997	1.0014	2.25	0.9777	2.1897
1.47	0.8061	1.0266	2.30	0.9806	2.2468
1.48	0.8122	1.0509	2.35	0.9831	2.3031
1.49	0.8180	1.0744	2.40	0.9853	2.3587
1.50	0.8235	1.0973	2.50	0.9890	2.4681
1.52	0.8339	1.1411	2.60	0.9918	2.5755
1.54	0.8436	1.1829	2.70	0.9939	2.6814
1.56	0.8525	1.2228	2.80	0.9955	2.7859
1.58	0.8608	1.2612	2.90	0.9967	2.8894
1.60	0.8685	1.2982	3.00	0.9976	2.9922

and

$$\sigma^2 = (\mu_f^2 + \sigma_f^2)/(1 + \theta^2), \tag{8}$$

where

$$G(\theta) = \frac{\sqrt{2/\pi}e^{-\theta^2/2} + \theta[1 - 2F(-\theta)]}{1 + \theta^2}, \qquad \theta = \mu/\sigma.$$

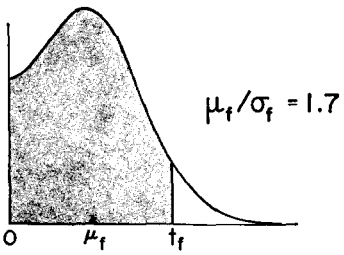
Table I gives values of σ_f and μ, both in σ units, corresponding to values of μ_f/σ_f so spaced that linear interpolation beyond $\mu_f/\sigma_f = 1.50$ yields values

TABLE II

Areas of the Folded Normal Distribution

(Classified According to μ_f/σ_f)

$$F(t_f) = \int_0^{t_f} f(x)dx \quad t_f \geq 0$$



t_f \ μ_f/σ_f	1.3236	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1
0.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.1	.0481	.0428	.0360	.0299	.0245	.0199	.0161	.0129	.0102
0.2	.0960	.0855	.0720	.0599	.0492	.0401	.0324	.0260	.0207
0.3	.1435	.1280	.1082	.0902	.0744	.0609	.0494	.0398	.0318
0.4	.1905	.1703	.1444	.1209	.1002	.0824	.0672	.0544	.0437
0.5	.2369	.2123	.1809	.1522	.1269	.1049	.0861	.0702	.0568
0.6	.2824	.2539	.2174	.1841	.1545	.1287	.1064	.0873	.0712
0.7	.3270	.2950	.2542	.2167	.1832	.1538	.1281	.1060	.0871
0.8	.3704	.3355	.2910	.2500	.2131	.1803	.1515	.1265	.1049
0.9	.4125	.3753	.3279	.2840	.2441	.2084	.1767	.1488	.1245
1.0	.4534	.4144	.3648	.3186	.2763	.2381	.2037	.1732	.1462
1.1	.4927	.4526	.4015	.3537	.3096	.2692	.2326	.1995	.1700
1.2	.5305	.4898	.4381	.3893	.3439	.3018	.2631	.2279	.1960
1.3	.5668	.5260	.4743	.4252	.3790	.3357	.2954	.2582	.2241
1.4	.6013	.5610	.5100	.4612	.4148	.3707	.3292	.2903	.2543
1.5	.6341	.5947	.5450	.4972	.4510	.4066	.3642	.3241	.2863
1.6	.6652	.6272	.5793	.5328	.4874	.4432	.4004	.3593	.3202
1.7	.6945	.6583	.6127	.5680	.5238	.4801	.4373	.3957	.3555
1.8	.7221	.6879	.6449	.6024	.5598	.5172	.4748	.4330	.3922
1.9	.7479	.7160	.6760	.6360	.5953	.5541	.5124	.4709	.4298
2.0	.7720	.7426	.7057	.6684	.6300	.5904	.5500	.5090	.4680
2.1	.7945	.7676	.7340	.6995	.6636	.6260	.5870	.5471	.5065
2.2	.8152	.7910	.7607	.7292	.6959	.6605	.6233	.5847	.5450
2.3	.8344	.8128	.7858	.7573	.7267	.6937	.6585	.6215	.5830
2.4	.8520	.8331	.8093	.7838	.7558	.7253	.6924	.6573	.6203
2.5	.8682	.8518	.8310	.8084	.7832	.7553	.7247	.6917	.6565
2.6	.8830	.8689	.8511	.8313	.8087	.7834	.7553	.7245	.6913
2.7	.8964	.8846	.8695	.8523	.8323	.8095	.7839	.7555	.7245
2.8	.9086	.8989	.8862	.8714	.8539	.8336	.8105	.7845	.7558
2.9	.9196	.9118	.9014	.8888	.8736	.8557	.8349	.8114	.7851
3.0	.9295	.9234	.9150	.9044	.8913	.8756	.8572	.8361	.8122
3.1	.9383	.9338	.9271	.9183	.9072	.8936	.8774	.8586	.8371
3.2	.9463	.9430	.9379	.9307	.9213	.9096	.8955	.8789	.8598
3.3	.9533	.9512	.9473	.9415	.9337	.9237	.9116	.8971	.8802
3.4	.9596	.9584	.9556	.9510	.9445	.9361	.9257	.9132	.8984
3.5	.9651	.9647	.9628	.9592	.9539	.9469	.9381	.9274	.9145

correct to approximately one in the fourth decimal place. One can thus make use of the estimates of μ_f and σ_f as given by $\bar{x}$ and s to estimate σ and μ from Table I.

3. Areas of the Folded Normal Distribution

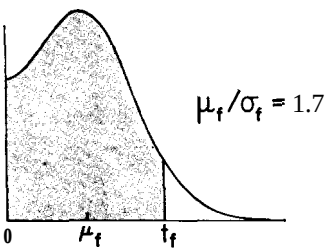
In many problems there may be no interest in relating the folded normal to a parent normal probability density function. One may wish to determine probabilities associated with the folded distribution. To solve such problems, Table II gives the areas under the folded normal from 0 to t_f for various values of μ_f/σ_f.

TABLE II (continued)

Areas of the Folded Normal Distribution

(Classified According to μ_f/σ_f)

$$F(t_f) = \int_0^{t_f} f(x)\,dx \quad t_f \text{ to}$$



t_f \ μ_f/σ_f	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.1	.0081	.0063	.0049	.0038	.0029	.0022	.0017	.0012	.0009
0.2	.0164	.0129	.0100	.0078	.0060	.0045	.0034	.0026	.0019
0.3	.0253						.0054	.0041	.0030
0.4	.0349	.0277	.0218	.0170	.0133	.0102	.0078	.0059	.0044
0.5	.0456	.0364	.0288	.0227	.0177	.0137	.0106	.0080	.0061
	.0576		.0370				.14		
0.7	.0711	.0577	.0464	.0371	.0231	.0180	.0180	.0107	.0082
0.8	.0863	.0706	.0573	.0462	.0370	.0029	.0232	.0140	.0108
	.1035							.0181	.0140
0.9	.1226	.1021	.0844	.0693	.0564	.0456	.0293	.0231	.0181
1.0							.0366	.0291	.0230
1.1	.1439	.1209	.1009	.0836	.0687	.0561	.0454	.0365	.0290
1.2	.1674	.1420	.1195				.0557	.0452	.0363
		.1653	.1404	.1184	.0991	.0823	.0679	.0555	.0450
1.3		.1909	.1637	.1392	.1176	.0985	.0819	.0676	.0553
1.4	.2512	.2188	.1892	.1624	.1383	.1169	.0981	.0816	.0674
1.5									
1.6	.2833	.2489	.2170	.1879	.1614	.1376	.1164	.0977	.0814
1.7	.3172						.1371		.0975
1.8	.3527	.3150	.2471	.2458	.1869	.1607	.1601	.1160	.1158
1.9	.3895	.3506	.3133	.2779	.2447	.2139	.1856	.1597	.1364
2.0	.4274	.3876	.3490	.3120	.2770	.2340	.2134	.1851	.1594
	.4659	.4256					.2434		.1848
2.2	.5047	.4643	.4243	.3850	.3141	.2762	.2757	.2129	.2126
2.3		.5034	.4632	.4233	.3838	.3104	.3099	.2753	.2427
2.4	.5435	.5424	.5025	4		3463		.3095	
2.5	.5819	.5811	.5417	.4.581	.4618	.423836	.3832	.3455	.3092
	.6560	.6189	.5805	.5411	.5013			4	
2.6	.6911	.6557	.6186	.5801	.5408	.4613	.4217	.3829	.3852
2.7			.6555		.5799	.5009	.5006	.4214	.3827
2.8	.7246	.6911	.6911	.6183	.6182	.5405	.5403	.4607	.4212
2.9	.7561	4		.6554		.5797			.4606
3.0	.7856	.7.7675	.7249	.6911	.6554	.6181	.5795	.5004	.5003
								.5401	
3.1	.8129	.7861	.7568	.7250	.6911	.6554	.6180	.5794	.5400
3.2	.8380	.8136	865		.7252	.6912		.6180	
3.3	.8607		.7865	.7570	.7572		.6554		.5794
3.4	.8812	.8615	.8393	.8145	.7872	.7253	.6554	.6913	.6554
3.5	.a995	.8820	.8622	.8398	.8148	.7874	.7576	.7255	.6913

It is interesting to observe that $F(t_f)$ for μ_f/σ_f equal to, say k, is not necessarily greater than $F(t_f)$ for μ_f/σ_f equal to $k' > k$. This is caused by μ_f and σ_f changing at different rates as the position of the fold changes.

4. An Illustration

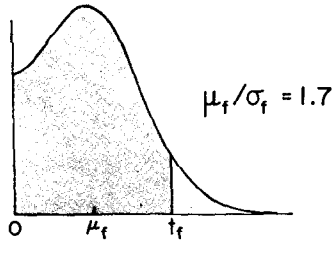
In the manufacture of miniature radio tubes, leads are fixed in place by automatic machinery. It is important that the leads be straight, otherwise, they may fail to drop into position and can jam the feeding mechanism. Straightness is measured on an optical comparator by clamping one end of the lead in a rotating chuck. The camber is reported as the differences between the maximum and

TABLE II (continued)

Areas of the Folded Normal Distribution

(Classified According to μ_f/σ_f)

$$F(t_f) = \int_0^{t_f} f(x)\,dx \quad t_f \geq 0$$



t_f \ μ_f/σ_f	1.3236	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1
3.6	.9700	.9702	.9690	.9663	.9620	.9562	.9488	.9397	.9287
3.7	.9743	.9750	.9744	.9723	.9688	.9641	.9580	.9503	.9409
3.8	.9780	.9791	.9789	.9774	.9747	.9708	.9657	.9593	.9514
3.9	.9813	.9826	.9828	.9816	.9795	.9764	.9723	.9670	.9604
4.0	.9841	.9856	.9860	.9852	.9836	.9811	.9778	.9735	.9680
4.1	.9865	.9882	.9887	.9882	.9869	.9850	.9823	.9788	.9743
4.2	.9887	.9903	.9909	.9906	.9897	.9882	.9860	.9832	.9796
4.3	.9905	.9921	.9927	.9926	.9919	.9907	.9891	.9868	.9839
4.4	.9920	.9936	.9942	.9942	.9937	.9928	.9915	.9897	.9874
4.5	.9933	.9948	.9955	.9955	.9951	.9945	.9935	.9921	.9902
4.6	.9944	.9959	.9964	.9965	.9963	.9958	.9950	.9939	.9925
4.7	.9954	.9967	.9972	.9973	.9972	.9968	.9962	.9954	.9943
4.8	.9961	.9974	.9979	.9980	.9978	.9976	.9971	.9965	.9957
4.9	.9969	.9979	.9983	.9984	.9984	.9982	.9979	.9974	.9968
5.0	.9974	.9984	.9987	.9988	.9988	.9987	.9984	.9981	.9976
5.1	.9979	.9987	.9990	.9991	.9991	.9990	.9988	.9986	.9982
5.2	.9983	.9990	.9993	.9994	.9993	.9993	.9992	.9990	.9987
5.3	.9986	.9992	.9995	.9995	.9995	.9995	.9994	.9993	.9991
5.4	.9989	.9994	.9996	.9997	.9997	.9996	.9996	.9995	.9993
5.5	.9991	.9995	.9997	.9997	.9998	.9997	.9997	.9996	.9995
5.6	.9993	.9997	.9998	.9998	.9998	.9998	.9998	.9997	.9997
5.7	.9994	.9997	.9998	.9999	.9999	.9999	.9998	.9998	.9998
5.8	.9995	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9998
5.9	.9996	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
6.0	.9997	.9999	.9999	1.0000	1.0000	1.0000	1.0000	.9999	.9999
6.1	.9998	.9999	1.0000					1.0000	.9999
6.2	.9998	.9999							1.0000
6.3	.9999	1.0000							
6.4	.9999								
6.5	.9999								
6.6	.9999								
6.7	.9999								
6.8	1.0000								
6.9									
7.0									

minimum positions assumed by the unclamped end upon rotation minus the diameter of the lead. Thus a lead that is perfectly straight will yield a camber of zero. It follows that camber is an inherently positive characteristic. Table III gives data taken from an experimental run.

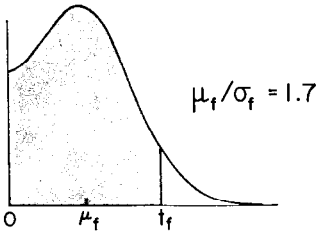
From Table III we can calculate $\bar{x} = 14.0140$, $s = 7.7868$, and $\bar{x}/s = 1.7997$. Using 1.80 as an estimate of μ_f/σ_f, Table I yields estimates of σ_f/σ and μ/σ equal to 0.9242 and 1.6192, respectively. This gives $\sigma = 8.4254$ and $\mu = 13.6424$. Substitution of these values in equations (1) and (2) gives the estimated theoretical normal and folded normal distributions, respectively. Figure 1 shows the data of Table III together with these distributions.

TABLE II (continued)

Areas of the Folded Normal Distribution

(Classified According to μ_f/σ_f)

$$F(t_f) = \int_0^{t_f} f(x)\,dx \quad t_f \geq 0$$



t_f \ μ_f/σ_f	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
3.6	.9156	.9003	.8827	.8627	.8401	.8151	.7876	.7577	.7255
3.7	.9297	.9164	.9010	.8832	.8631	.8404	.8153	.7877	.7578
3.8	.9419	.9305	.9171	.9015	.8836	.8634	.8407	.8155	.7878
3.9	.9524	.9427	.9311	.9176	.9019	.8840	.8636	.8408	.8156
4.0	.9613	.9531	.9433	.9316	.9180	.9022	.8842	.8638	.8410
4.1	.9688	.9619	.9536	.9437	.9320	.9183	.9025	.8844	.8639
4.2	.9750	.9694	.9624	.9541	.9441	.9323	.9185	.9027	.8845
4.3	.9802	.9756	.9698	.9628	.9544	.9444	.9325	.9187	.9028
4.4	.9844	.9807	.9760	.9702	.9631	.9547	.9446	.9327	.9189
4.5	.9879	.9848	.9810	.9763	.9705	.9634	.9549	.9447	.9328
4.6	.9906	.9882	.9851	.9813	.9765	.9707	.9636	.9550	.9449
4.7	.9928	.9909	.9885	.9854	.9815	.9767	.9708	.9637	.9551
4.8	.9945	.9931	.9911	.9887	.9856	.9817	.9769	.9710	.9638
4.9	.9959	.9947	.9932	.9913	.9888	.9857	.9818	.9770	.9710
5.0	.9969	.9960	.9949	.9934	.9914	.9889	.9858	.9819	.9770
5.1	.9977	.9971	.9962	.9950	.9935	.9915	.9890	.9859	.9820
5.2	.9983	.9978	.9972	.9963	.9951	.9936	.9916	.9891	.9859
5.3	.9988	.9984	.9979	.9972	.9963	.9952	.9936	.9917	.9891
5.4	.9991	.9988	.9985	.9980	.9973	.9964	.9952	.9937	.9917
5.5	.9994	.9992	.9989	.9985	.9980	.9973	.9964	.9952	.9937
5.6	.9996	.9994	.9992	.9989	.9986	.9980	.9974	.9965	.9953
5.7	.9997	.9996	.9994	.9992	.9990	.9986	.9981	.9974	.9965
5.8	.9998	.9997	.9996	.9995	.9993	.9990	.9986	.9981	.9974
5.9	.9999	.9998	.9997	.9996	.9995	.9993	.9990	.9986	.9981
6.0	.9999	.9999	.9998	.9997	.9996	.9995	.9993	.9990	.9986
6.1	.9999	.9999	.9999	.9998	.9997	.9996	.9995	.9993	.9990
6.2	1.0000	.9999	.9999	.9999	.9998	.9998	.9996	.9995	.9993
6.3		1.0000	.9999	.9999	.9999	.9998	.9998	.9997	.9995
6.4			1.0000	.9999	.9999	.9999	.9998	.9998	.9997
6.5				1.0000	.9999	.9999	.9999	.9998	.9998
6.6					1.0000	.9999	.9999	.9999	.9998
6.7						1.0000	.9999	.9999	.9999
6.8							1.0000	.9999	.9999
6.9								1.0000	.9999
7.0									1.0000

TABLE 3

Camber of 497 Lead Wires

Camber*	Observed Frequency	Camber*	Observed Frequency
1	12	23	18
3	26	25	14
5	36	27	12
7	50	29	4
9	45	31	7
11	49	33	2
13	51	35	5
15	44	37	2
17	44	39	2
19	40	41	1
21	32	43	1

* In mils. Class mid-points are given.

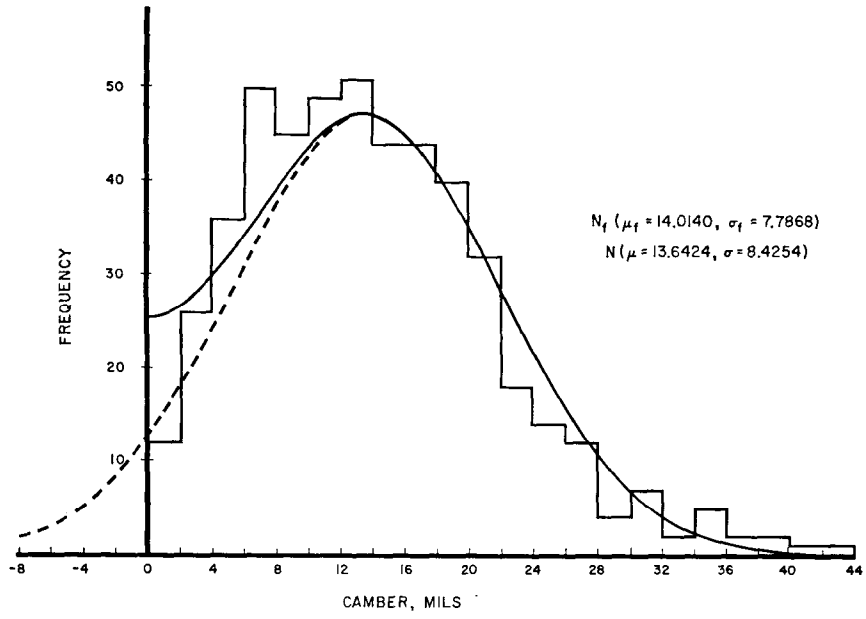


FIGURE 1 – Camber of 497 lead wires.

ACKNOWLEDGMENT

We wish to acknowledge the assistance given by F. W. Kuhlman in the computation of Tables I and II.

REFERENCE

1. DANIEL, CUTHBERT, "Use of Half-Normal Plots in Interpreting Two Level Experiments", *Technometrics, 1*, 311–341 (1959).