JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 76B, Nos. 3 and 4, July–December 1972

Confidence Limits for the Abscissa of Intersection of Two Linear Regressions

James J. Filliben

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

and

John E. McKinney

Institute for Materials Research, National Bureau of Standards, Washington, D.C. 20234

(August 25, 1972)

A method to determine the confidence limits for the abscissa of the intersection of two linear regressions has been developed. This method does not require the assumption of equal variance for the two regressions, as was necessary with previous methods. A numerical example is included on thermodynamic, glass transition data for which this method is applicable. Comparisons are made between the results using equal and unequal variance assumptions. A FORTRAN subroutine is included for computations using both assumptions.

Key words: Abscissa; confidence limits; glass transition; intercept; intersection; linear; regression; second order transition; statistics; variance.

1. Introduction

In this paper confidence limits are derived for the abscissa of the intersection point of two linear regressions. As illustrated in figure 1, it is assumed that there are independent realizations of two linear regressions at ordered values of a single independent variable X. The problem is to determine confidence limits for the abscissa X_0 of the intersection point of the two linear regressions.

This problem was originally considered and solved by Fisher [1][1] and later by Kastenbaum [2], but under the assumption of a common error variance for the two regression lines. Hinkley [3] derived the asymptotic distribution of the maximum likelihood estimator for the intersection point abscissa of two linear regressions, also assuming equal variances. Robison [4] presented estimators and confidence limits for the intersection point abscissa of two polynomial regressions under the equal variance assumption. Hudson [5] derived least squares estimators (but no confidence limits) for the intersection point abscissas of two or more general regression models for the unequal variance case.

The principal contribution of the present paper is a derivation of confidence limits for the intersection point abscissa without resorting to the equal variance assumption. Although only the linear regression case is considered below, extension of results to more complicated regression models is possible.

The results of this derivation have direct application to thermodynamic data to evaluate the uncertainties in the second order transition temperature T_2 and the glass transition temperature T_g.

AMS Subject Classification: 62.55.

[1] Figures in brackets indicate the literature references at the end of this paper.

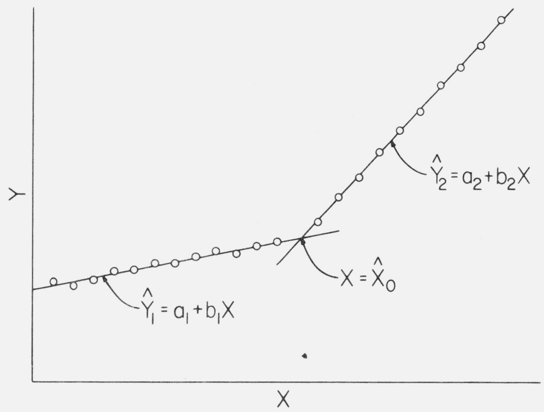

FIGURE 1. *Intersection of two linear regressions.*

The second order transition is an equilibrium phenomenon, whereas the glass transition is a non-equilibrium one for which T_g depends upon the thermodynamic history by which the glass is formed. Both transitions may be manifested by discontinuities in the second order properties (for example, thermal expansivity or heat capacity) with respect to temperature. The corresponding first order properties (volume or enthalpy) vary with temperature in a manner similar to Y versus X in figure 1 for which the transition temperature may be determined by the intersection of the two straight lines representative of the response in the two regions. Accordingly, this method may be applied directly to any of the first order properties to determine the confidence bands of the transition temperature as defined[2] from the intersection of the two regression lines. The method is sufficiently general that it is applicable to cases where data are not available in the proximity of the intersection, which, of course, leads to a less accurate prediction of its abscissa.

Two methods are described to evaluate the confidence limits for the abscissa of intersection. These are based on the assumptions of equal and unequal variances for the regression lines determined from the two sets. The advantage of the equal variance case is that the underlying assumption of equal variance permits a relatively simple development of the method, whereas the advantage of the unequal variance case is, of course, in its more general application. The ANSI FORTRAN subroutine to calculate the confidence limits for both cases is given in the appendix, and numerical examples are included in later discussions.

2. Derivation of Confidence Limits

As illustrated in figure 1, let $\hat{Y}_1(X) = a_1 + b_1X$ and $\hat{Y}_2(X) = a_2 + b_2X$ be two fitted least squares regressions based on n_1 and n_2 observations (X_{ij}, $i=1, 2; j=1, 2, \ldots, n_i$) respectively, where the random variables $\hat{Y}_i(X)$ are assumed to be normally and independently distributed about the true unknown regression lines $\mu_i(X) = \alpha_i + \beta_i X$ with variances σ_i^2. (α_i, β_i, and σ_i^2 are true, but unknown, parameter values.)

The problem is to determine confidence limits for the abscissa X_0 of the intersection point (X_0, Y_0) of the two unknown regression lines.

The starting point for our derivation is the fact that a_i and b_i, the usual least squares estimators, are linear combinations of normally distributed variables and hence themselves are normally distributed. Thus for all X, $\hat{Y}_i(X) = a_i + b_iX$ is also normally distributed, as is the difference $\hat{Y}_1(X) - \hat{Y}_2(X)$. Since $\hat{Y}_1(X) - \hat{Y}_2(X)$ has mean $\mu_1(X) - \mu_2(X)$ and variance

[2] This definition of the glass temperature is one of several which are not necessarily equivalent.

$\sigma_1^2 V_1(X) + \sigma_2^2 V_2(X)$, where

$$V_i(X) = \frac{1}{n_i} + \frac{(X - \bar{X}_i)^2}{S_i^2}, \qquad i = 1, 2$$

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, \qquad i = 1, 2 \tag{1}$$

$$S_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2, \qquad i = 1, 2,$$

then it follows that all X

$$\frac{[\hat{Y}_1(X) - \hat{Y}_2(X)] - [\mu_1(X) - \mu_2(X)]}{\sqrt{\sigma_1^2 V_1(X) + \sigma_2^2 V_2(X)}} \tag{2}$$

has a standardized (mean 0, variance 1) normal distribution.

Consider first of all the equal variance case ($\sigma_1^2 = \sigma_2^2 = \sigma^2$, say), treated earlier [1]. It is seen that eq (2) simplifies to

$$\frac{[\hat{Y}_1(X) - \hat{Y}_2(X)] - [\mu_1(X) - \mu_2(X)]}{\sigma \sqrt{V_1(X) + V_2(X)}}. \tag{3}$$

Thus

$$\frac{[\hat{Y}_1(X) - \hat{Y}_2(X)] - [\mu_1(X) - \mu_2(X)]}{s \sqrt{V_1(X) + V_2(X)}} \tag{4}$$

has a t distribution with $n_1 + n_2 - 4$ degrees of freedom, where s is the square root of the pooled unbiased estimator of σ^2:

$$s = \sqrt{\frac{(n_1 - 2)s_1^2 + (n_2 - 2)s_2^2}{n_1 + n_2 - 4}}, \tag{5}$$

and where s_i^2 is the unbiased estimator for σ_i^2 given by the residual variance obtained from the least squares fit of line i.

The quantity (4) is distributed as t for all X, and in particular for $X = X_0$ in which case $\mu_1(X) - \mu_2(X)$ vanishes. It then follows that for every probability $p (0 < p < 1)$

$$\text{Prob}\left\{\left(\frac{\hat{Y}_1(X_0) - \hat{Y}_2(X_0)}{s\sqrt{V_1(X_0) + V_2(X_0)}}\right)^2 \leqslant \left(G_t\left(\frac{1+p}{2}\right)\right)^2\right\} = p \tag{6}$$

where $G_t\left(\frac{1+p}{2}\right)$ is the $100\left(\frac{1+p}{2}\right)$ percent point of the t distribution with $n_1 + n_2 - 4$ degrees of freedom. The squared numerator $[\hat{Y}_1(X_0) - \hat{Y}_2(X_0)]^2$ and squared denominator $s^2[V_1(X_0) + V_2(X_0)]$ above are both quadratics in X_0. Thus eq (6) is of the form

$\text{Prob}\left\{\frac{Q_1}{Q_2} \leqslant G^2\right\}$ (or equivalently Prob $\{Q_1 - G^2 Q_2 \leqslant 0\}$) where Q_i denote quadratic expressions, and G denotes $G_t\left(\frac{1+p}{2}\right)$. Since $Q_1 - G^2 Q_2$ is itself a quadratic (say $AX_0^2 + BX_0 + C$), then the desired confidence interval consists of those X_0 such that $AX_0^2 + BX_0 + C \leq 0$ and the desired confidence limits for X_0 are simply the two roots of $AX_0^2 + BX_0 + C = 0$.

The coefficients A, B, and C are rather tediously derived; omitting the intermediate algebra, these coefficients are

$$A = (b_1 - b_2)^2 - s^2G^2\left[\frac{1}{S_1^2} + \frac{1}{S_2^2}\right]$$

$$B = 2(a_1 - a_2)(b_1 - b_2) + 2s^2G^2\left[\frac{\bar{X}_1}{S_1^2} + \frac{\bar{X}_2}{S_2^2}\right] \tag{7}$$

$$C = (a_1 - a_2)^2 - s^2G^2\left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{\bar{X}_1^2}{S_1^2} + \frac{\bar{X}_2^2}{S_2^2}\right].$$

It is of interest to note that the midpoint $-B/2A$ of the confidence interval is not identical to the maximum likelihood estimator

$$\hat{X}_0 = -\frac{a_1 - a_2}{b_1 - b_2}, \tag{8}$$

however, the two estimators converge to a common limit as $s^2 \to 0$. For a less detailed derivation of the above confidence interval results, see references [1–3].

For the more general $\sigma_1^2 \neq \sigma_2^2$ case, the distributional complications which arise are identical to those encountered in the Behrens-Fisher 2-means problem [6]. The solution outlined below is an extension of the Welch-Aspin [7, 8, 9] solution to the Behrens-Fisher problem. Proceeding in a manner analogous to Welch and Aspin, we see that for all X,

$$\frac{[\hat{Y}_1(X) - \hat{Y}_2(X)] - [\mu_1(X) - \mu_2(X)]}{\sqrt{s_1^2V_1(X) + s_2^2V_2(X)}} \tag{9}$$

has approximately a t distribution with ν degrees of freedom, where ν is given by

$$\frac{1}{\nu} = \frac{1}{n_1-2}\left(\frac{V_1(X_0)}{V_1(X_0)+V_2(X_0)}\right)^2 + \frac{1}{n_2-2}\left(\frac{V_2(X_0)}{V_1(X_0)+V_2(X_0)}\right)^2 \tag{10}$$

A minor complication at this point is the fact this X_0 in (10) is an unknown parameter. In order to arrive at an approximate value for ν the maximum likelihood estimate $\hat{X}_0$ (as defined in (8)) is used. The development now parallels the $\sigma_1^2 = \sigma_2^2$ case, and so again 100 p confidence limits for X_0 are obtained by solving a quadratic equation, but with the following slightly-modified coefficients:

$$A = (b_1 - b_2)^2 - G^2\left[\frac{s_1^2}{S_1^2} + \frac{s_2^2}{S_2^2}\right]$$

$$B = 2(a_1 - a_2)\ (b_1 - b_2) + 2G^2\left[\frac{s_1^2\bar{X}_1}{S_1^2} + \frac{s_2^2\bar{X}_2}{S_2^2}\right] \tag{11}$$

$$C = (a_1 - a_2)^2 - G^2\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} + \frac{s_1^2\bar{X}_1^2}{S_1^2} + \frac{s_2^2\bar{X}_2^2}{S_2^2}\right]$$

Note that although the unequal variance case basically parallels the equal variance case, the former does not numerically reduce to the latter when $s_1^2 = s_2^2$. The confidence intervals usually are distinct even when $s_1^2 = s_2^2$. In deciding whether to use the equal variance results or the unequal variance results, there are two considerations to be taken into account: physical theory and the variance ratio F test. The general rule of thumb is to always use the unequal variance results unless the following two conditions are simultaneously satisfied: (1) there is a firm physical reason for assuming equal variances; and (2) the F ratio S_1^2/S_2^2 is not significantly different from unity. This recommended procedure is at times conservative, but more important, it is less assumption-dependent and hence more generally valid.

In the appendix, an ANSI FORTRAN subroutine is presented which first of all tests to see whether the equal variance assumption is tenable, and secondly computes confidence limits for X_0 for both the equal variance and the unequal variance cases. The subroutine is double precision in input, internal operation, and output. Definitions of input parameters and instructions for the use of the subroutine are given in the comment statements within the subroutine.

3. Example: Calculation of the 95 percent Confidence Limits for the Pressure-Dependent Glass Temperature

As an example, the 95 percent confidence limits have been calculated for the pressure-dependent glass transition temperature of poly(vinyl acetate), an amorphous, glass-forming polymer. The PVT (pressure-volume-temperature) data for the glass, designated as set 1, are given in table 1, and the data for the liquid, designated as set 2, are given in table 2. Both sets of data,

TABLE 1. *Specific volume values for the glassy region*

T/P	0	100	200	300	400	500	600	700	800
−30	0.82884	0.82654	0.82438	0.82225	0.82017	0.81816	0.81613	0.81419	0.81237
−25	.83002	.82768	.82547	.82334	.82121	.81914	.81715	.81516	.81327
−20	.83110	.82874	.82648	.82430	.82218	.82010	.81808	.81606	.81415
−15	.83238	.82997	.82767	.82549	.82331	.82117	.81919	.81717	.81522
−10	.83350	.83117	.82886	.82663	.82443	.82224	.82022	.81817	.81620
−5	.83467	.83233	.83003	.82774	.82557	.82340	.82131	.81933	.81729
0	.83581	.83351	.83115	.82889	.82666	.82447	.82237	.82033	.81825
5	.83706	.83478	.83241	.83012	.82794	.82569	.82355	.82140	.81941
10	.83812	.83593	.83359	.83130	.82903	.82679	.82465	.82248	.82040
15	.83937								
20	.84061								

Units: T in °C, P in Bars, v in cm^3/g.

TABLE 2. *Specific volume data for the liquid region*

T/P	0	100	200	300	400	500	600	700	800
35	0.84572	0.84148							
40	.84870	.84444	0.84041	0.83670					
45	.85174	.84733	.84321	.83921	0.83549	0.83184			
50	.85486	.85041	.84608	.84206	.83817	.83447	0.83087	0.82757	
55	.85791	.85349	.84894	.84491	.84092	.83709	.83340	.82988	0.82666
60	.86104	.85628	.85179	.84769	.84367	.83980	.83607	.83256	.82904
65	.86407	.85933	.85472	.85052	.84648	.84250	.83874	.83523	.83174
70	.86723	.86218	.85755	.85324	.84913	.84507	.84108	.83734	.83378
75	.87038	.86536	.86038	.85594	.85171	.84762	.84367	.84000	.83626
80	.87343	.86829	.86342	.85881	.85453	.85036	.84641	.84257	.83889
85	.87669	.87140	.86636	.86169	.85728	.85308	.84911	.84532	.84155
90	.87986	.87438	.86923	.86444	.86000	.85574	.85172	.84776	.84396
95	.88301	.87741	.87207	.86722	.86269	.85823	.85424	.85025	.84641
100	.88622	.88043	.87500	.87012	.86534	.86095	.85678	.85284	.84904

Units: T in °C, P in bars, v in cm^3/g taken from reference [10], are illustrated in figure 2, where the specific volume v taken from the above tables is plotted as a function of temperature at the different pressures indicated. The specific volume data $v_1(T, P)$ for the glass (at temperatures between -30 and 20 °C) were obtained after the glass was formed at atmospheric pressure at a constant cooling rate of 5 °C/h commencing at equilibrium in the liquid region. The data $v_2(T, P)$ for the liquid (35 to 100 °C) are taken to be in true equilibrium. A characteristic of a glass is its large increase in its viscoelastic relaxation times with decreasing temperature. The glass transition is a manifestation of the increase of relaxation times during formation of the glass. Although the data shown here for the glass are not representative of the attainment of true equilibrium, the relaxation times at these temperatures are very long in comparison to experimental times. Accordingly, the glass data may be treated as being in apparent equilibrium (pseudoequilibrium) over any practical time scale, as has been done in this case.

Since the volume isobars vary nearly linearly with temperature in both the liquid and glass regions, and since the slopes for each isobar in these regions are distinct, the results of section 2 are applicable. By treating each of the two regions separately, each of the volume isobars may be represented by two distinct regression lines of the form

$$\hat{v}_i = a_i + b_i T, \qquad i = 1, 2. \tag{12}$$

The intersection of these two regression lines at each pressure gives the "maximum likelihood estimator"

$$\hat{T}_g = -\frac{a_1 - a_2}{b_1 - b_2}, \tag{13}$$

which is taken here as an estimate of the true, but unknown, glass temperature T_g for the particular thermodynamic history by which the glass was formed. In figure 2 the regression lines of the form of eq (12) are shown by the solid lines with positive slope at each pressure. The solid line

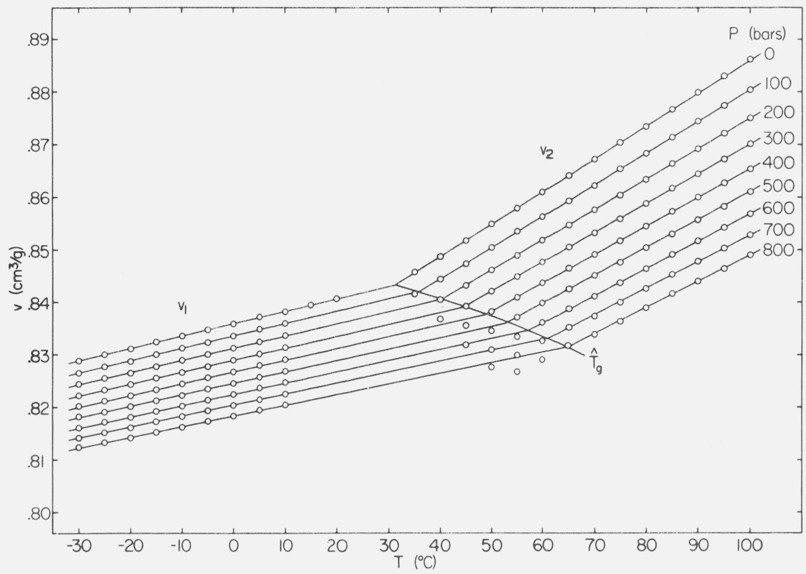

FIGURE 2. *Specific volume of poly(vinyl acetate) as a function of temperature at different pressures.*

An estimate of the glass transition is obtained from the intersection of the volume isobars taken to be linear with temperature.

with negative slope gives the positions of $\hat{T}_g$, which depend strongly on the pressure. It is seen that some of the data representative of the liquid region fall to the left of $\hat{T}_g$. Although this result may be paradoxical, it is simply a consequence of the fact that the liquid data are equilibrium data, and the glass are pseudoequilibrium data[4] resulting from the particular thermodynamic history by which the glass was formed.

Ninety-five percent confidence limits for T_g were calculated using both the equal and unequal variance assumptions. The paired statistical quantities necessary to calculate the confidence limits are given in tables 3 and 4 for the glass and liquid, respectively, where P is the pressure, a_i is the intercept, b_i is the slope, s_i is the standard deviation of the residuals about the regression lines,

TABLE 3. *Statistical data for the Glassy Region*

P	a_1	$b_1 \times 10^3$	$s_1 \times 10^4$	n_1	$\bar{T}_1$	S_1^2
0	0.835843	0.23469	0.495	11	−5.0	2750.0
100	.833531	.23590	.565	9	−10.0	1500.0
200	.831192	.23123	.778	9	−10.0	1500.0
300	.828928	.22660	.855	9	−10.0	1500.0
400	.826716	.22270	.910	9	−10.0	1500.0
500	.824512	.21713	1.021	9	−10.0	1500.0
600	.822417	.21327	.793	9	−10.0	1500.0
700	.820328	.20860	.737	9	−10.0	1500.0
800	.818301	.20270	.897	9	−10.0	1500.0

TABLE 4. *Statistical data for the Liquid Region*

P	a_2	$b_2 \times 10^3$	$s_2 \times 10^4$	n_2	$\bar{T}_2$	S_2^2
0	0.823686	0.62355	1.146	14	67.5	5687.5
100	.820417	.59934	.869	14	67.5	5687.5
200	.817195	.57760	.683	13	70.0	4550.0
300	.814189	.55834	.832	13	70.0	4550.0
400	.811039	.54362	.565	12	72.5	3575.0
500	.808010	.52946	.589	12	72.5	3575.0
600	.804836	.51987	.952	11	75.0	2750.0
700	.802096	.50698	1.341	11	75.0	2750.0
800	.799213	.49703	1.564	10	77.5	2062.5

n_i is the number of observations, $\bar{T}_i$ is the average temperature, and S_i^2 is the sum of the squared temperature deviations:

$$S_i^2 = \sum_{j=1}^{n_i} (T_{ij} - \bar{T}_i)^2, \qquad i = 1, 2.$$

The results of these calculations are summarized in table 5 for both the equal and unequal variance cases. G is the 97.5 percent point of the t_ν distribution with ν degrees of freedom where $\nu = n_1 + n_2 - 4$ in the equal variance case, and where ν is defined by eq (10) in the unequal variance case. $-A/2B$ is the midpoint of the confidence interval where A and B are defined by eqs (7) in the equal variance case and eqs (11) in the unequal variance case. ΔT is the 95 percent confidence interval width which is equal to the difference of the two roots of the quadratic equation for which the constants are defined in eqs (7) and (11) for equal variance and unequal variance, respectively. s_1/s_2 is the ratio of residual standard deviations, and R is the ratio of ΔT for the equal variance case to ΔT for the unequal variance case. Perusal of table 5 reveals several interesting points. First, it is noted that the corresponding midpoints $-B/2A$ are essentially identical for the equal

[4] In principle, at least at temperatures over a limited range below T_g shown here, the specific volume values for the glass will relax to extrapolations of the liquid isobars shown provided sufficient time is allowed. Since this time may be as long as several million years (depending on the temperature and pressure), the attainment of true equilibrium in all cases (as with the liquid values shown) is not practical.

TABLE 5. *Estimated values of* T_g *and 95 percent Confidence Intervals*

P	$\hat{T}_g$	Equal variance				Unequal variance				$\frac{s_1}{s_2}$	R
		ν	G	$-B/2A$	ΔT	ν	G	$-B/2A$	ΔT		
0	31.26	21	2.08	31.27	0.92	17	2.11	31.26	0.79	0.432	1.162
100	36.08	19	2.09	36.09	1.18	9	2.26	36.09	1.02	0.650	1.158
200	40.41	18	2.10	40.42	1.26	9	2.26	40.42	1.44	1.138	0.874
300	44.43	18	2.10	44.44	1.62	8	2.31	44.44	1.80	1.028	0.898
400	48.85	17	2.11	48.86	1.56	8	2.31	48.87	2.07	1.609	0.752
500	52.84	17	2.11	52.85	1.84	8	2.31	52.86	2.53	1.734	0.729
600	57.34	16	2.12	57.36	2.24	8	2.31	57.35	2.21	.833	1.014
700	61.10	16	2.12	61.13	3.03	7	2.36	61.11	2.34	.549	1.296
800	64.85	15	2.13	64.89	3.77	7	2.36	64.87	3.02	.573	1.248

and unequal variance cases, and these, in turn, are essentially identical to the maximum likelihood estimator $\hat{T}_g$. As mentioned earlier these quantities converge to the common limit $\hat{T}_g$ as $s_i^2 \to 0$. Secondly, the confidence limits vary considerably between the equal and unequal variance cases (the ratio R being at times both greater than or less than unity). In the case where s_1 and s_2 are nearly equal ($P = 300$ bars), R differs considerably from unity; and in the case where R is nearly unity ($P = 700$ bars), s_1 and s_2 are considerably distinct. This behavior illustrates the fact that the equal variance case is not a special case of the unequal variance case as might be expected. Thirdly, there appears to be little correlation between the ratio of confidence interval widths and the ratio of residual standard deviations.

The values of the maximum likelihood estimator $\hat{T}_g$ and those for the 95 percent confidence interval width ΔT (each centered about its midpoint $-B/2A$) for both the equal variance ($\sigma_1^2 = \sigma_2^2$) and the unequal variance ($\sigma_1^2 \neq \sigma_2^2$) cases are shown on figure 3 as functions of pressure. The median solid line is a quadratic regression of $\hat{T}_g$ with respect to pressure. The two extreme solid lines are quadratic regressions of the confidence limits with respect to pressure for the unequal variance case only. It is clear that a trend exists between the confidence interval and pressure. The principal reason for the divergence of the confidence interval with increasing pressure is that the difference $(\hat{T}_g - \bar{T}_1)$ between the estimated glass temperature $\hat{T}_g$ and the average temperature for the glass $\bar{T}_1$ also increases considerably with pressure as seen from figure 2. Thus is apparent the importance of obtaining data in the vicinity of the glass temperature (or, in general, the intersection) in order to obtain an accurate estimate.

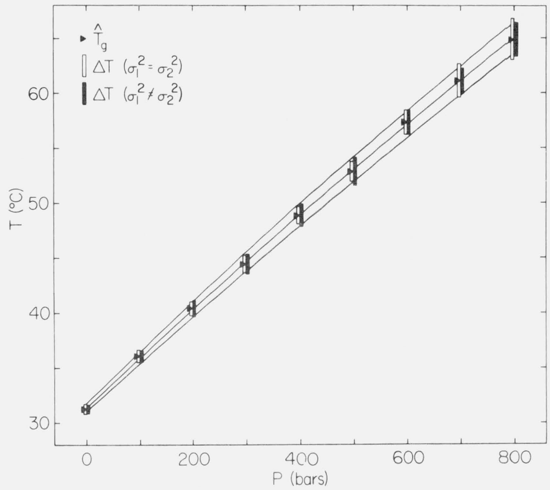

FIGURE 3. *Maximum Likelihood Estimator,* $\hat{T}_g$, *and 95 percent Confidence Intervals as functions of pressure.*

The solid lines are quadratic regressions of these quantities.

4. Appendix

```
      SUBROUTINE ABSCIS(AN1,AN2,XBAR1,XBAR2,SSQDX1,SSQDX2,ALP1,ALP2,BET1
     1,BET2,RSD1,RSD2)
C     THE PURPOSE OF THIS ROUTINE IS TO COMPUTE 95 PERCENT CONFIDENCE LIMITS
C     FOR THE ABSCISSA OF THE INTERSECTION POINT OF TWO LINEAR REGRESSIONS
C     FOR BOTH THE EQUAL VARIANCES CASE AND THE UNEQUAL VARIANCES CASE
C     AN1 = THE NUMBER OF (X,Y) PAIRS FOR LINE 1
C     AN2 = THE NUMBER OF (X,Y) PAIRS FOR LINE 2
C     XBAR1 = THE SAMPLE MEAN OF THE X VALUES FOR LINE 1
C     XBAR2 = THE SAMPLE MEAN OF THE X VALUES FOR LINE 2
C     SSQDX1 = THE SUM OF THE SQUARED DEVIATIONS OF THE X VALUES FOR LINE 1
C              ABOUT XBAR1
C     SSQDX2 = THE SUM OF THE SQUARED DEVIATIONS OF THE X VALUES FOR LINE 2
C              ABOUT XBAR2
C     ALP1 = THE Y INTERCEPT OF LINE 1
C     ALP2 = THE Y INTERCEPT OF LINE 2
C     BET1 = THE SLOPE OF LINE 1
C     BET2 = THE SLOPE OF LINE 2
C     RSD1 = THE RESIDUAL STANDARD DEVIATION OF LINE 1
C     RSD2 = THE RESIDUAL STANDARD DEVIATION OF LINE 2
C     THIS ROUTINE IS DOUBLE PRECISION IN INPUT, INTERNAL OPERATION, AND OUTPUT
      DIMENSION Z(30)
      DOUBLE PRECISION AN1,AN2,XBAR1,XBAR2,SSQDX1,SSQDX2
      DOUBLE PRECISION ALP1,ALP2,BET1,BET2,RSD1,RSD2
      DOUBLE PRECISION Z
      DOUBLE PRECISION P,S,DELALP,DELBET,AML
      DOUBLE PRECISION ANUEQ,ANUNEQ,GEQ,GNEQ,PP,HPP
      DOUBLE PRECISION RATIO,AM,AN,HOLD
      DOUBLE PRECISON A1,A2,A3,A,B1,B2,B3,B,C1,C2,C3,C
      DOUBLE PRECISION DISC,ROOT1,ROOT2,AMID,WIDTH1,WIDTH2,RATWID
      DOUBLE PRECISION DSQRT
      DATA Z(1),Z(2),Z(3),Z(4),Z(5),Z(6),Z(7),Z(8),Z(9),Z(10),Z(11),
     1Z(12),Z(13),Z(14),Z(15),Z(16),Z(17),Z(18),Z(19),Z(20),Z(21),
     1Z(22),Z(23),Z(24),Z(25),Z(26),Z(27),Z(28),Z(29),Z(30)/12.7026D0,
     14.3027D0,3.1824D0,2.7764D0,2.5706D0,12.4469D0,2.3646D0,2.3060D0,
     12.2622D0,2.2281D0,2.2010D0,2.1788D0,12.1604D0,2.1448D0,2.1315D0,
     12.1199D0,2.1096D0,2.1009D0,2.0930D0,12.0860D0,2.0796D0,2.0739D0,
     12.0687D0,2.0639D0,2.0595D0,2.0555D0,12.0518D0,2.0484D0,2.0452D0,
     12.0423D0/
C
C     COMPUTE NEEDED CONSTANTS
C
      P=.95D0
      S=DSQRT(((AN1-2.0D0)*RSD1*RSD1+(AN2-2.0D0)*RSD2*RSD2)/(AN1+AN2-4.0
     1D0))
      DELALP=ALP1-ALP2
      DELBET=BET1-BET2
C
C     COMPUTE THE MAXIMUM LIKELIHOOD ESTIMATOR
0
      AML=-DELALP/DELBET
C
C     COMPUTE THE APPROPRIATE DEGREES OF FREEDOM FOR THE T PERCENT POINTS
C     FOR THE EQUAL VARIANCES CASE AND FOR THE UNEQUAL VARIANCES CASE
C
      ANUEQ=AN1+AN2-4.0D0
      V1=(1.0D0/AN1)+(AML-XBAR1)*(AML-XBAR1)/SSQDX1
      V2=(1.0D0/AN2)+(AML-XBAR2)*(AML-XBAR2)/SSQDX2
      ANUNEQ=(1.0D0/(AN1-2.0D0))*((V1/(V1+V2))**2)+(1.0D0/(AN2-2.0D0))*(
     1(V2/(V1+V2))**2)
      ANUNEQ=1.0D0/ANUNEQ
```

```
      NUEQ=ANUEQ+0.1D0
      NUNEQ=ANUNEQ+0.1D0
      IF(NUEQ.LE.30)GEQ=Z(NUEQ)
      IF(NUEQ.GT.30)GEQ=1.96D0*DSQRT(ANUEQ/(ANUEQ-2.0D0))
      IF(NUNEQ.LE.30)GNEQ=Z(NUNEQ)
      IF(NUNEQ.GT.30)GNEQ=1.96D0*DSQRT(ANUNEQ/(ANUNEQ-2.0D0))
      PP=(P+1.0D0)/2.0D0
      HPP=100.0D0*PP
C
C     WRITE OUT THE INPUT DATA
C     IT IS ASSUMED THAT THE OUTPUT UNIT IS 6.  IF THIS IS NOT THE CASE, THEN
C     CHANGE THE NEXT STATEMENT APPROPRIATELY.
C
      IP=6
      WRITE(IP,998)
      WRITE(IP,170)
      WRITE(IP,999)
      WRITE(IP,180)
      WRITE(IP,999)
      WRITE(IP,190)AN1,AN2
      WRITE(IP,200)XBAR1,XBAR2
      WRITE(IP,210)SSQDX1,SSQDX2
      WRITE(IP,220)ALP1,ALP2
      WRITE(IP,230)BET1,BET2
      WRITE(IP,240)RSD1,RSD2
      WRITE(IP,999)
      WRITE(IP,250)P
      WRITE(IP,999)
      WRITE(IP,260)HPP,NUEQ,GEQ
      WRITE(IP,270)HPP,NUNEQ,GNEQ
C
C     COMPUTE THE RATIO OF THE LARGER ESTIMATED RESIDUAL VARIANCE TO
C     THE SMALLER ESTIMATED RESIDUAL VARIANCE
C
      RATIO=(RSD1*RSD1)/(RSD2*RSD2)
      AM=AN1-2.0D0
      AN=AN2-2.0D0
      IF(RATIO.GT.1.0D0)GOTO110
      RATIO=1.0D0/RATIO
      HOLD=AM
      AM=AN
      AN=HOLD
110   WRITE(IP,999)
      M1=AM+0.1
      N1=AN+0.1
      WRITE(IP,280)RATIO
      WRITE(IP,310)M1,N1
      WRITE(IP,999)
      WRITE(IP,290)AML
      WRITE(IP,999)
      WRITE(IP,300)S
      WRITE(IP,999)
      WRITE(IP,999)
C
C     COMPUTE THE COEFFICIENTS OF THE QUADRATIC EQUATION FOR THE EQUAL
C     VARIANCES CASE
C
      WRITE(IP,340)
      WRITE(IP,999)
      A1=DELBET**2
      A2=S*S*GEQ*GEQ
```

```
      A3=(1.0D0/SSQDX1)+(1.0D0/SSQDX2)
      B1=2.0D0*DELALP*DELBET
      B2=2.0D0*A2
      B3=(XBAR1/SSQDX1)+(XBAR2/SSQDX2)
      C1=DELALP**2
      C2=A2
      C3=(1.0D0/AN1)+(1.0D0/AN2)+(XBAR1*XBAR1/SSQDX1)+(XBAR2*XBAR2/SSQDX
     12)
      A=A1-A2*A3
      B=B1+B2*B3
      C=C1-C2*C3
C
C     CHECK FOR ZERO COEFFICIENT OF THE QUADRATIC TERM FOR THE EQUAL
C     VARIANCES CASE
C
      IF(A.NE.0.0D0)GOTO120
      AMID=-C/B
      WRITE(IP,320)
      WRITE(IP,321)AMID
      GOTO140
C
C     COMPUTE THE ROOTS OF THE QUADRATIC, AND THE MIDPOINT OF THE ROOTS
C     FOR THE EQUAL VARIANCES CASE
C
  120 DISC=B*B-4.0D0*A*C
C
C     CHECK FOR A NEGATIVE DISCRIMINANT FOR THE EQUAL VARIANCES CASE
C
      IF(DISC.GE.0.0D0)GOTO130
      WRITE(IP,330)DISC
      GOTO140
  130 ROOT1=(-B+DSQRT(DISC))/(2.0D0*A)
      ROOT2=(-B-DSQRT(DISC))/(2.0D0*A)
      AMID=-B/(2.0D0*A)
      WIDTH1=ROOT1-ROOT2
      WRITE(IP,350)ROOT2
      WRITE(IP,360)ROOT1
      WRITE(IP,370)AMID
      WRITE(IP,380)WIDTH1
      WRITE(IP,999)
      WRITE(IP,999)
C
C     COMPUTE THE COEFFICIENTS OF THE QUADRATIC EQUATION FOR THE UNEQUAL
C     VARIANCES CASE
C
140   WRITE(IP,420)
      WRITE(IP,999)
      A1=DELBET**2
      A2=GNEQ*GNEQ
      A3=(RSD1*RSD1/SSQDX1)+(RSD2*RSD2/SSQDX2)
      B1=2.0D0*DELALP*DELBET
      B2=2.0D0*A2
      B3=(RSD1*RSD1*XBAR1/SSQDX1)+(RSD2*RSD2*XBAR2/SSQDX2)
      C1=DELALP**2
      C2=A2
      C3=(RSD1*RSD1/AN1)+(RSD2*RSD2/AN2)+(RSD1*RSD1*XBAR1*XBAR1/SSQDX1)+
     1(RSD2*RSD2*XBAR2*XBAR2/SSQDX2)
      A=A1-A2*A3
      B=B1+B2*B3
      C=C1-C2*C3
C     CHECK FOR ZERO COEFFICIENT OF THE QUADRATIC TERM FOR THE UNEQUAL
```

```
C     VARIANCES CASE
      IF(A.NE.0.0D0)GOTO150
      AMID=-C/B
      WRITE(IP,400)
      WRITE(IP,321)AMID
      RETURN
C
C     COMPUTE THE ROOTS OF THE QUADRATIC, AND THE MIDPOINT OF THE ROOTS
C     FOR THE UNEQUAL VARIANCES CASE
C
  150 DISC=B*B-4.0D0*A*C
C
C     CHECK FOR A NEGATIVE DISCRIMINANT FOR THE UNEQUAL VARIANCES CASE
C
      IF(DISC.GE.0.0D0)GOTO160
      WRITE(IP,410)DISC
      RETURN
  160 ROOT1=(-B+DSQRT(DISC))/(2.0D0*A)
      ROOT2=(-B-DSQRT(DISC))/(2.0D0*A)
      AMID=-B/(2.0D0*A)
      WIDTH2=ROOT1-ROOT2
      RATWID=WIDTH2/WIDTH1
      WRITE(IP,350)ROOT2
      WRITE(IP,360)ROOT1
      WRITE(IP,370)AMID
      WRITE(IP,380)WIDTH2
      WRITE(IP,999)
      WRITE(IP,999)
      WRITE(IP,390)RATWID
      RETURN
C
C
  170 FORMAT(1H ,58HSTATISTICAL ANALYSIS OF INTERSECTION OF 2 REGRESSION
     1 LINES)
  180 FORMAT(1H ,10HINPUT DATA)
  190 FORMAT(1H ,9HAN1     = ,D15.8,10X,9HAN2     = ,D15.8)
  200 FORMAT(1H ,9HXBAR1   = ,D15.8,10X,9HXBAR2   = ,D15.8)
  210 FORMAT(1H ,9HSSQDX1 = ,D15.8,10X,9HSSQDX2 = ,D15.8)
  220 FORMAT(1H ,9HALP1    = ,D15.8,10X,9HALP2    = ,D15.8)
  230 FORMAT(1H ,9HBET1    = ,D15.8,10X,9HBET2    = ,D15.8)
  240 FORMAT(1H ,9HRSD1    = ,D15.8,10X,9HRSD2    = ,D15.8)
  250 FORMAT(1H ,52H2-SIDED CONFIDENCE INTERVALS WITH PROBABILITY WIDTH
     1,D12.5,13H ARE PROVIDED)
  260 FORMAT(1H ,4HTHE ,D12.5,42H PERCENT POINT OF THE T DISTRIBUTION WI
     1TH ,I6,50H DEGREES OF FREEDOM (THE EQUAL VARIANCES CASE) IS ,D12.5
     1)
  270 FORMAT(1H ,4HTHE ,D12.5,42H PERCENT POINT OF THE T DISTRIBUTION WI
     1TH ,I6,52H DEGREES OF FREEDOM (THE UNEQUAL VARIANCES CASE) IS ,D12
     1.5)
  280 FORMAT(1H ,98HTHE RATIO OF THE LARGER ESTIMATED RESIDUAL VARIANCE
     1TO THE SMALLER ESTIMATED RESIDUAL VARIANCE IS ,D15.8)
  290 FORMAT(1H ,35HTHE MAXIMUM LIKELIHOOD ESTIMATOR = ,D15.8)
  300 FORMAT(1H ,4HS = ,D15.8)
  310 FORMAT(1H ,79HCOMPARE THE ABOVE RATIO VALUE TO THE PERCENT POINTS
     1OF THE F DISTRIBUTION WITH ,I6,5H AND ,I6,19H DEGREES OF FREEDOM)
  320 FORMAT(1H ,86H*****ERROR--THE COEFFICIENT OF THE QUADRATIC TERM IS
     1 ZERO FOR THE EQUAL VARIANCES CASE)
  321 FORMAT(1H ,57H               THE CONFIDENCE INTERVAL HAS ZERO WIDTH A
     1BOUT ,D15.8)
  330 FORMAT(1H ,97H*****ERROR--NEGATIVE DISCRIMINANT IN THE EQUAL VARIA
     1NCES CASE.  THE VALUE OF THE DISCRIMINANT IS ,D15.8)
```

```
  340 FORMAT(1H ,20HEQUAL VARIANCES CASE)
  350 FORMAT(1H ,37HLOWER LIMIT OF CONFIDENCE INTERVAL = ,D15.8)
  360 FORMAT(1H ,37HUPPER LIMIT OF CONFIDENCE INTERVAL = ,D15.8)
  370 FORMAT(1H ,37HMIDPOINT OF CONFIDENCE INTERVAL    = ,D15.8)
  380 FORMAT(1H ,37HWIDTH OF CONFIDENCE INTERVAL       = ,D15.8)
  390 FORMAT(1H ,46HTHE RATIO OF THE CONFIDENCE INTERVAL WIDTHS = ,D15.8
     1)
  400 FORMAT(1H ,88H*****ERROR--THE COEFFICIENT OF THE QUADRATIC TERM IS
     1 ZERO FOR THE UNEQUAL VARIANCES CASE)
  410 FORMAT(1H ,99H*****ERROR--NEGATIVE DISCRIMINANT IN THE UNEQUAL VAR
     1IANCES CASE.  THE VALUE OF THE DISCRIMINANT IS ,D15.8)
  420 FORMAT(1H,22HUNEQUAL VARIANCES CASE)
  998 FORMAT(1H1)
  999 FORMAT(1H0)
      END
```

```
STATISTICAL ANALYSIS OF INTERSECTION OF 2 REGRESSION LINES


INPUT DATA


AN1    =   .14000000+002           AN2    =    .11000000+002
XBAR1  =   .67500000+002           XBAR2  =  -.50000000+001
SSQDX1 =   .56875000+004           SSQDX2 =    .27500000+004
ALP1   =   .82368597+000           ALP2   =    .83584345+000
BET1   =   .62355168-003           BET2   =    .23469094-003
RSD1   =   .11459900-003           RSD2   =    .49539027-004


2-SIDED CONFIDENCE INTERVALS WITH PROBABILITY WIDTH   .95000+000 ARE PROVIDED


THE  .97500+002 PERCENT POINT OF THE T DISTRIBUTION WITH   21 DEGREES OF
     FREEDOM (THE EQUAL VARIANCES CASE) IS  .20796+001
THE  .97500+002 PERCENT POINT OF THE T DISTRIBUTION WITH   17 DEGREES OF
     FREEDOM (THE UNEQUAL VARIANCES CASE) IS  .21098+001


THE RATIO OF THE LARGER ESTIMATED RESIDUAL VARIANCE TO THE SMALLER ESTIMATED
     RESIDUAL VARIANCE IS  .5351393+001
COMPARE THE ABOVE RATIO VALUE TO THE PERCENT POINTS OF THE F DISTRIBUTION
     WITH  .12000+002 AND  .90000+001 DEGREES OF FREEDOM


THE MAXIMUM LIKELIHOOD ESTIMATOR =   .31264354+002


S =   .92500246-004
```


479-206 O - 73 - 7

```
EQUAL VARIANCES CASE

LOWER LIMIT OF CONFIDENCE INTERVAL =    .30804202+002
UPPER LIMIT OF CONFIDENCE INTERVAL =    .31727843+002
MIDPOINT OF CONFIDENCE INTERVAL    =    .31266023+002
WIDTH OF CONFIDENCE INTERVAL       =    .92364138+000


UNEQUAL VARIANCES CASE

LOWER LIMIT OF CONFIDENCE INTERVAL =    .30865357+002
UPPER LIMIT OF CONFIDENCE INTERVAL =    .31660331+002
MIDPOINT OF CONFIDENCE INTERVAL    =    .31262844+002
WIDTH OF CONFIDENCE INTERVAL       =    .79497346+000


THE RATIO OF THE CONFIDENCE INTERVAL WIDTHS =   .86069493+000
```

5. References

[1] Fisher, R. A., Statistical Methods for Research Workers, 11th edition, 144–146 (Oliver and Boyd, 1950).

[2] Kastenbaum, M. A., A confidence interval on the abscissa of the point of intersection of two fitted linear regressions, Biometrics **15**, 323–324 (1959).

[3] Hinkley, D. V., Inference about the intersection in two-phase regression, Biometrika **56**, 495–504 (1969).

[4] Robison, D. E., Estimate for the points of intersection of two polynomial regressions, J. Am. Statistical Asn. **59**, 214–224 (1964).

[5] Hudson, D. J., Fitting segmented curves whose join points have to be estimated, J. Amer. Stat. Assoc. **61**, 1097–1129 (1966).

[6] Bennett, Carl A. and Franklin, Norman L., Statistical Analysis in Chemistry and the Chemistry Industry (John Wiley & Sons, New York, 1954). pp 176–180.

[7] Aspin, A. A., An examination and further development of a formula arising in the problem of comparing two mean values, Biometrika **35**, 88 (1948).

[8] Aspin, A. A., Tables for use in comparisons whose accuracy involves two variances, separately estimated, Biometrika **36**, 290 (1949).

[9] Welch, B. L., The generalisation of "Student's" problem when several different population variances are involved, Biometrika **34**, 28 (1947).

[10] McKinney, J. E., and Belcher, H. V., to be published.

(Paper 76B3&4–370)