貧血検査の結果の統計学的考察

唐 戸 哲 哉

埼玉県衛生研究所

(昭和60年9月30日 受付)

1 はじめに

貧血検査の目的で普通ヘモグロビン値の測定が行われる。金井¹⁾によれば、ヘモグロビンの定量法には、種々の方法があり臨床的には、従来ザリー法が用いられたが、近年シアンメトヘモグロビン法が、普及している。ザリー法は、臨床上には、簡便であるが、誤差はかなり大きい。最も正確な方法は Van Slyke の検圧計を用いて、血液の酸素結合能、または、一酸化炭素結合能を測定し、あるいは血液の鉄含量を測定して、これより、Hb 量を算出する方法である。これに次いで、シアンメトヘモグロビン法による、光電光度法であって、国際血液学標準委員会 (ICSH) は、本法によることを推奨している。全血及び血漿の比重から (硫酸銅法) 計算する方法もあり、同時に血漿蛋白濃度を知ることが出来る。

次に、血球と血漿の容積比は、正常人では一定しているが、貧血では、その程度に応じて減少する。したがって、ヘマトクリット値 (Ht 値) は貧血の尺度となり、赤血球数あるいは、ヘモグロビン濃度の測定と大体同様の意味を有し、かつ測定法が、簡便であり、また正確度が高い。Ht 値と赤血球数から、平均赤血球容積 MCV が求められ、貧血の種類の鑑別に利用することができる。ヘマトクリット値の測定と同時に、白血球数の概数、血漿の色調や透明度などの、参考所見が得られる。検査法として、ヘマトクリット法、屈折計法、粘稠度計法、放射能同位元素法、比重により測定する方法など種々あるが、臨床上には、ヘマトクリット法が、簡便で、最もよく用いられる。

私は、結核住民検診の際に行われた、貧血検査の結果得た、ヘモグロビン値、ヘマトクリット値の結果を、統計学的に考察したので、その結果について、報告する。

II 測定方法

血色素量 (Hb と略す) の測定—シアンメトヘモグロビン法を用いた。血球容積 (Haematocrit 値, Ht と略す) の測定—マイクロヘマトクリット法を用いた。

III 結 果

A. Hb 値の度数分布について

10月18日、採血液の Hb 測定値の受付番号の 1 番から 50 番まで選んで、これを任意標本と考えて正規分布しているか否かについて考えた。その度数分布図を図 1 に記したが、正規曲線の適合検査を表 1 のようにして行なった。結果の χ^2 の値は、表のように、14.01、あるいは 6.07 で自由度 df $11 - 3 = 8$, $df = 8 - 3 = 5$ で有意の差を示さなかった。 ($P \leq 0.05$, あるいは $P \leq 0.01$) 又表 2 のようにして、NED²⁾ を計算して、プロットした。結果は、図 2 に示すように、直線化されず、正規性を証明することは、出来なかった。

B. 男女間の Hb 値の比較について

男女間の Hb 値を比較するために、任意標本 10 づつを選んで、平均値を比較してみた。その結果は、男性では 14.5g/dl、女性では、12.4g/dl で、男性の方が高かった。t の値は、3.25 で自由度 $= 10 + 10 - 2 = 18$ で有意の差を示した。 ($P \leq 0.01$) また、男女間の 2 標本平均値の差をみるために、 $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$ の仮説をたて、CR を計算してみた。

$$CR = \frac{X_1 - X_2 - \mu X_1 - \mu X_2}{\sigma X_1 - X_2} =$$

$$\frac{14.3 - 12.4}{\sqrt{\frac{1.09}{10}} + \sqrt{\frac{2.63}{10}}} = \frac{1.9}{\sqrt{\frac{3.72}{10}}} =$$

$$\frac{1.9}{\sqrt{0.372}} = \frac{1.9}{0.61} = 3.11$$

図1 ヘモグロビンの度数分布 $n=50$ $\bar{x}=12.19$ $S=1.72$ $V=2.95$ (g/dL)

表1 度数分布の適合度の χ^2 検定

x_i	f_i	$x_i - \bar{x}$	$\frac{x_i - \bar{x}}{S} = Z$	$n \times \phi(Z)$	Σf_{ei}	te_i	$f_i - fe$	$\frac{(f_i - fe)^2}{fe_i}$
5.5—6.0	1	4	7					
6.0—6.5	0							
6.5—7.0	0							
7.0—7.5	0							
7.5—8.0	0	7	7					
8.0—8.5	3							
9.0—9.5	0							
9.5—10.0	0							
10.0—10.5	2	-2.19	-1.27	-19.89	5.11	5.11	-1.11	0.24
10.5—11.0	1	-1.69	-0.98	-16.82	8.18	3.07	-1.07	0.37
11.0—11.5	5	-1.19	-0.69	-12.74	12.26	4.08	-3.08	2.32
11.5—12.0	6	-0.69	-0.40	-7.77	17.23	4.97	-0.03	0.00
12.0—12.5	7	-0.19	-0.11	-1.99	23.01	5.78	0.22	0.00
12.5—13.0	8	+0.31	+0.18	+3.57	28.57	5.56	1.44	0.37
13.0—13.5	6	+0.81	+0.47	+9.04	34.04	5.47	2.53	1.17
13.5—14.0	6	+1.31	+0.76	+13.81	38.81	4.77	1.23	0.31
14.0—14.5	6	+1.81	+1.05	+17.65	42.65	3.84	2.16	1.21
14.5—	5	+2.31	+1.34	+20.49	45.49	2.84	3.16	3.51
計	50					50		14.01

$n=50$ $\bar{x}=12.19$ $S=1.72$ $V=2.95$

$af=11-3=8$ $\chi^2_{0.05}(8)=15.51 > \chi^2$ $df=8-3=5$ $\chi^2_{0.05}(5)=11.07 > \chi^2$
 $\chi^2_{0.01}(8)=20.1 > \chi^2$ $\chi^2_{0.01}(5)=15.09 > \chi^2$

図2 血色素分布の正規性の検討

表2 N.E.D. の計算

ヘモグロビン値 (g/dl)	被 検 数 (N)	累 積 度 数	累 積 度 数 $\Sigma(f) - 0.5(\%)$	N. E. D.
5.5—6.0	1	1	0.01	-2.36
6.0—6.5	0	1	0.01	-2.36
6.5—7.0	0	1	0.01	-2.36
7.0—7.5	0	1	0.01	-2.36
7.5—8.0	0	1	0.01	-2.36
8.0—8.5	3	4	0.07	-1.476
8.5—9.0	0	4	0.07	-1.476
9.0—9.5	0	4	0.07	-1.476
9.5—10.0	0	4	0.07	-1.476
10.0—10.5	2	6	0.11	-1.227
10.5—11.0	1	7	0.13	-1.126
11.0—11.5	5	12	0.23	-0.739
11.5—12.0	6	18	0.35	-0.385
12.0—12.5	7	25	0.49	-0.025
12.5—13.0	8	33	0.65	+0.385
13.0—13.5	6	39	0.77	+0.739
13.5—14.0	6	45	0.89	+1.227
14.0—14.5	5	50	0.99	+2.326

表3 男女間のヘモグロビン値の比較

	男	女
Σ	143.3	124.4
n	10	10
$\bar{x}$	14.3	12.4
$S.D.$	1.23	1.67
t -値	3.25	

表4 各年代におけるヘモグロビン値の比較

	20才代	30才代	40才代	50才代	60才代
Σ	66.2	62.3	61.7	62.8	63.7
f	5	5	5	5	5
$\bar{x}$	13.2	12.4	12.3	12.5	12.7
F -値	F_{20}^4 0.87				

この程度の差の生ずる確率は、 $P \leq 0.01 \therefore \mu_1 = \mu_2$ の仮説をすて、2標本平均値間に有意の差があると判定した。

C. 年齢階級による、Hb値の平均値の差について
20代、30代、40代と10代毎に、Hb値の平均値を出してみた。(n=5)その結果は、表4にあるように、40代、30代、50代、60代、20代の順に平均値が高くなっている。 $F_{20}^4 = 0.87$ で各平均値間に有意の差を示さなかった。(P ≥ 0.05)

D. 母集団平均値 μ の推定について

Aに書いたように、標本度数分布は、 χ^2 検定では正規分布を示したが、母集団も正規分布をするものと考えて、95%と99%の母集団平均値 μ の信頼区間を、前記の集団 (n=50, s (σ)=1.72, $\bar{x}$ =12.19) を用いて計算した。

95%の信頼区間

$$\bar{x} - 1.96 \frac{S}{\sqrt{N}} < \mu < \bar{x} + 1.96 \frac{S}{\sqrt{N}}$$

$$12.19 - 1.96 \frac{1.72}{\sqrt{50}} < \mu < 12.19 + 1.96 \frac{1.72}{\sqrt{50}}$$

$$11.72 < \mu < 12.66$$

99%の信頼区間

$$\bar{x} - 2.58 \frac{S}{\sqrt{N}} < \mu < \bar{x} + 2.58 \frac{S}{\sqrt{N}}$$

$$\bar{x} - 2.58 \frac{1.72}{\sqrt{50}} < \mu < \bar{x} + 2.58 \frac{1.72}{\sqrt{50}}$$

$$12.19 - 2.58 \frac{1.72}{\sqrt{50}} < \mu < 12.19 + 2.58 \frac{1.72}{\sqrt{50}}$$

$$12.19 - 2.58 \frac{1.72}{7.07} < \mu < 12.19 + 2.58 \frac{1.72}{7.07}$$

$$12.19 - 2.58 \times 0.243 < \mu < 12.19 + 2.58 \times 0.243$$

$$12.19 - 0.626 < \mu < 12.19 + 0.626$$

$$11.56 < \mu < 12.81$$

E. 正規分布の分散の信頼区間

Dのように、母集団が、正規分布に近い分布をするものと、考えて、分散の95%の信頼区間を求めてみた。標本として、受付番号1から、10までのHb値を用いた。

$$\Sigma \chi^2 = 124.4 \quad \Sigma \chi^2 = 1573.8 \quad n = 10$$

$$\Sigma (x_i - \bar{x})^2 = 1573.88 - 1547.53$$

$$= 26.35$$

$$\frac{26.35}{19.02} < \sigma^2 < \frac{26.35}{2.70}$$

$$1.38 < \sigma^2 < 9.75$$

F. Hb値、Ht値の直線回帰、相関、回帰係数の信頼区間、有意性の検定について

まづ、Hb値をY値(従属変数)、Ht値を、x値(独立変数)として考察した。Hb値、Ht値間の回帰方程式を、受付番号の1から10までの値を用いて求めた。その式は、 $y = 2.55x + 6.18$ という式だった。永田³⁾によれば、一般に回帰方程式は、 $y_a = \mu + \beta(x_a - \bar{x}) + \epsilon_a$ という式をうる。実測の個々の y_a の値を $\bar{y}$ (μ の推定値)、 $b(x_a - \bar{x})$ 、 $\beta(x_a - \bar{x})$ の推定値、 $e(\epsilon_a$ の推定値)の3部分に分割すると、表5を得る。ただし表5では、 e は標本回帰方程式、 $y_a = 37.8 + 2.55(x_a - 12.4) + ed$ から求められた、 βa の値と実測値 y_a との差として求められている。

$$S^2_{yx} = \Sigma d^2 yx / 10 - 2 = 4.38 / 8 = 0.548$$

bの有意性の検定

$$F = \frac{b^2 \Sigma (x - \bar{x})^2}{S^2_{yx}} = \frac{184.55}{0.548} = 336.77$$

したがって、有意の差あり (P ≤ 0.01)

$$\text{又上式は、} t^2 = \frac{b^2}{S^2_{yx} / \Sigma (x - \bar{x})^2}$$

$$t = \frac{b}{\sqrt{S^2_{yx} / \Sigma (x - \bar{x})^2}}$$

$$= \frac{b}{S_{yx} / \sqrt{\Sigma (x - \bar{x})^2}}$$

$$S_b = \frac{S_{yx}}{\sqrt{\Sigma (x - \bar{x})^2}} = \frac{\sqrt{4.38}}{\sqrt{26.36}}$$

$$= \frac{2.09}{5.13} \doteq 0.407$$

表5 個々の実測値 y_2 の分割

No.	x	y_2	x'	y'	$b(x-\bar{x})$	e	e^2
1	13.4	41	1.0	3.2	2.55	0.65	0.423
2	12.2	37	-0.2	-0.8	-0.51	-0.29	0.084
3	12.5	39	0.1	1.2	0.25	0.945	0.893
4	8.4	27	-4.0	-10.8	-10.20	-0.6	0.360
5	13.6	40.5	1.2	2.7	3.06	-0.36	0.129
6	12.0	37	-0.4	-0.8	-1.02	0.22	0.048
7	14.4	41.5	2.0	3.7	5.10	-1.4	1.960
8	14.1	42	1.7	4.2	4.33	-0.13	0.016
9	11.5	35	-0.9	-2.8	-2.29	-0.51	0.2601
10	12.3	38	-0.1	0.2	-0.25	0.455	0.2070
平均	12.4	37.8					$\Sigma 4.38$

$$S_{xx}=26.36 \quad S_{yy}=176.1 \quad S_{xy}=67.28 \quad F_{0.05}(1.8)=5.32$$

$$b=S_{xy}/S_{xx}=2.55 \quad SE=176.1-2.55 \times 67.28=4.54$$

$$F=(10-2) \times 2.55 \times 67.28 / 4.536=302.6$$

$$t = \frac{b}{S_b} = \frac{2.55}{0.407} = 6.26$$

自由度 $10-2=8$ の t 分布で、1%有意水準の t の値は、3.355 5%有意水準では、2.306で何れよりも t の値は大きい。したがって、 b は 1%、5%有意水準で、有意であると判定される。

b の標本標準偏差、 S_b を用いると、母回帰係数 β の信頼区間を求めることができる。

β の 95%の信頼区間は、

$$b-t_{0.05} S_b \leq \beta \leq b+t_{0.05} S_b$$

自由度 8 の $P=0.05$ の t の値は、2.306

$$S_b=0.407, \quad b=2.55$$

$$2.55-2.306 \times 0.407 \leq \beta \leq 2.55+2.306 \times 0.407$$

$$2.55-0.938 \leq \beta \leq 2.55+0.938$$

$$1.612 \leq \beta \leq 3.488$$

又 β の 99%の信頼区間は、

自由度 $=n-2=10-2=8$ $P=0.01$ のときの t の値は 3.355なので

$$2.55-3.355 \times 0.407 \leq \beta \leq 2.55+3.355 \times 0.407$$

$$2.55-1.365 \leq \beta \leq 2.55+1.365$$

$$1.185 \leq \beta \leq 3.915$$

G. Hb 値と Ht 値の標本相関係数、母相関係数

前記の F の Hb 値は y 値(従属変数)、Ht 値を、 x 値(独立変数)として、考えたのであるが、本項では 2 変数間の関連性という点から考えてみる。

まづ、受付番号の、1 番から 10 番(女性)の、Hb

値、Ht 値から $\gamma=0.98$ という、標本相関係数を得た。この値を使って、 $\rho=0$ の母相関係数を検定してみた。その結果は $t=1.75$ で、 $Df=10-2=8$ の t の値は $P=0.05$ のときには、2.306、 $P=0.01$ のときには、3.355 で帰無仮説 $\rho=0$ は否定されなかった。

次に、 $P=99\%$ 、95%の信頼区間を求めてみた。 $\gamma=0.98$ は、 $Z=2.51$ 、 $n=10$ であるから、 $\sigma^2 = \frac{1}{n-3}$

$$= \frac{1}{7} \quad Df=\infty \text{ のときの } t \text{ 分布の有意水準, } P=0.01$$

のとき、2.576

$$2.51-2.576 \times 0.378 \leq Z \leq 2.51+2.51 \times 0.378$$

$$1.536 \leq Z \leq 3.4838$$

$$0.910 \leq \rho \leq 3.4838$$

又 95%の信頼区間は

$$0.944 \leq \rho \leq 0.999$$

次に、男女間の相関の大きさを、比較するために、男性から任意標本 10 をとり ($n=10$) $\gamma=0.94$ の値を得て、女性の $\gamma=0.98$ の値と比較してみた。 $\gamma=0.94$ は、 $Z=1.7381$ 、 $\gamma=0.98$ は、 $Z=2.2976$ 、この Z の

$$\text{差 } Z_d=0.5595 \text{ は分散 } S^2_z = \frac{1}{n_1-3} + \frac{1}{n_2-3}$$

$$= \frac{1}{7} + \frac{1}{7} = \frac{2}{7} = 0.2857 \text{ であり, } Df=\infty \text{ の}$$

正規分布をするとみなしうる。

$$t=Z_d / \sqrt{S^2_z}$$

$$=0.5595/\sqrt{0.2858}$$

$$=0.5595/0.5346$$

$$=1.046$$

Df = ∞ の時の, $P=0.01$ の値は, 2.576であり, したがって有意ではない.

H. Hb 値と Ht 値の順位相関について, Hb 値と Ht 値を性別に, 小さい値から番号をつけ, 次式によって γ_s の値を求めてみた. ($n=8$)

$$\gamma_s = 1 - \frac{6 (\sum di^2)}{N (N^2 - 1)}$$

その結果, 男性では $\gamma_s=0.75$, 女性では $\gamma_s=0.95$ で, 男性では, $P \leq 0.05$ のとき, 女性では $P \leq 0.01$ のときに有意だった.

IV 結 論

貧血検査の目的で, 測定された, Hb 値, Ht 値を統計学的に, 考察することによって次のような結論に達した.

A. 任意標本 ($n=50$) を選んで Hb 値の度数分布を χ^2 検査, NED の直線化で調べたところ χ^2 検査では正規性を証明したが, NED は直線化されなかった.

B. 男女間の Hb の値を平均値で, 比較してみると, 男性が高かった. 又 t の値, CR の値でみると, $P \leq 0.01$ で有意の差を示した.

C. 年齢階級により, Hb の値を平均値でみると, 40代, 30代, 50代, 60代, 20代の順に高かったが, F_{20}^4 の値は有意の差を示さなかった. ($P \geq 0.05$)

D. 母集団平均値 μ の 95%, 99% の信頼区間を, 前記の集団 ($n=50$, $s=1.72$, $\bar{x}=12.19$) を用いて計算した.

E. 正規分布に近い分布の分散の信頼区間を求めた.

$$1.38 < \sigma^2 < 9.75$$

F. Hb 値, Ht 値の直線回帰, 相関, 回帰係数の信頼区間, 有意性について, 検討を加えた. 回帰方程式は, $y=2.55x+6.18$ という式であり, 自由度 $10-2=8$ で, b は 1%, 5% の有意水準で有意ではないと, 判定された. β の 95%, 99% の信頼区間は次のようだった.

$$1.612 \leq \beta \leq 3.488$$

$$1.185 \leq \beta \leq 3.915$$

G. Hb 値と Ht 値の標本相関係数, 母相関係数, 有意性と大きさの比較を求めてみた. その結果, $\gamma=0.98$ という, $n=10$ の女性の相関係数は, $\rho=0$ の母集団から, しばしば出るものであることが証明された.

又 ρ の 99%, 95% の信頼区間は次のようである.

$$0.910 \leq \rho \leq 3.483$$

$$0.944 \leq \rho \leq 0.999$$

次に男性から, $n=10$ の標本をとり, $\gamma=0.94$ の値を得て, 大きさを比較したが, 有意ではなかった.

H. Hb 値と Ht 値の順位相関を性別に求めた.

この研究をするにあたって, 埼玉県衛生研究所病理細菌部疫学部の方々の御援助をうけました. 又埼玉県公害センター 伍井邦宏科長の御援助をうけました. 厚くお礼申し上げます. 又, 衛生研究所長河内卓先生の御校閲に深謝します.

文 献

- 1) 金井 泉, 金井正光: 臨床検査法提要第29版, 金原出版, 東京 1983, 235頁.
- 2) 辻 達彦: 統計方法入門 第3版. 金原出版, 東京 1977, 179-187頁.
- 3) 永田久紀: 医学公衆衛生学のための統計学入門. 南江堂. 東京 1972, 64-72頁.

STATISTICAL STUDIES ON ANEMIA**TETSUYA KARATO**

Saitama Institute of Public Health

The correlation between hemoglobin concentration and hematocrit value was studied, as follows. 1) By χ^2 -test and NED-test, hemoglobin concentrations of random samples were determined, and hemoglobin concentrations showed a normal distribution with the former test, but not with the latter. 2) Average values of hemoglobin concentrations in men and women were compared by T-test, and Cr-test. 3) The arithmetic means of hemoglobin concentrations for each ten-year age group and F values were tested. 4) Confidence intervals for averages of population values were estimated. 5) Confidence limits of variance of nearly normal distribution were estimated. 6) Linear regression, correlation, confidence interval of regression coefficient and significance for hemoglobin concentration and hematocrit value were tested. 7) Coefficients of correlation between sample and population were compared as to their significance. 8) Spearman's rank correlation was examined.

Key words : Hematocrit, Hemoglobin, Anemia, Statistical study.