# Vergleich der Meßgenauigkeit von Rauchmaschinen\*

von H. Riedwyl

Institut für Mathematische Statistik und Versicherungslehre der Universität Bern, Schweiz

#### 1. EINLEITUNG

Zur Zeit ist eine ganze Serie verschiedener automatischer Maschinen für das Abrauchen von Cigaretten in Betrieb. In den USA ist als Standardmaschine die Rauchmaschine von *Philip Morris* (3) eingeführt worden, doch kam es für die übrigen Länder im allgemeinen nicht zu einer Standardisierung.

Lipp (2) hat eine Übersicht über einige Standard-Rauchmaschinen gegeben und die Rauchausbeute verschiedener Maschinen an Hand einer Stichprobe verglichen. Da nicht alle Maschinen die gleiche Anzahl von Cigaretten nacheinander oder simultan abrauchen, erschwert sich der statistische Vergleich der Meßgenauigkeit verschiedener Maschinen.

Im folgenden soll gezeigt werden, was wir unter Meßgenauigkeit verstehen wollen und wie die einzelnen Komponenten statistisch beurteilt werden können. Dabei stößt man insbesondere auf einen verallgemeinerten t-Test, der weit über dieses Anwendungsgebiet hinaus Verwendung finden kann.

## 2. MESSGENAUIGKEIT

Das wiederholte Abrauchen einer festen Anzahl Cigaretten mit einer Rauchmaschine nach einem gut definierten Meßverfahren ergibt Meßwerte (z. B. Feuchtkondensat, Trockenkondensat, Rauchnikotin, etc.), die mehr oder weniger voneinander abweichen und deren Mittelwerte selbst von Maschine zu Maschine oder von einem bestimmten Referenzwert abweichen.

Unter der *Präzision* wollen wir den zufälligen Fehler einer Rauchmaschine verstehen. Werden unter möglichst gleichen Umweltbedingungen Messungen gemacht, so kann die Einzelmessung als Zufallsvariable mit bekanntem oder unbekanntem Verteilungsgesetz betrachtet werden. Ist der Mittelwert dieser Zufallsvariable  $\mu$ , so kann als Maß der Präzision eine Funktion der Abweichungen der Einzelmessungen x von diesem Mittelwert  $\mu$  oder  $\bar{x}$  angesehen werden. In der Regel zeichnet man die Standardabweichung (oder deren Quadrat, die Varianz),

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} , \qquad [1]$$

als Schätzwert der theoretischen Varianz σ aus.

Das Ausmaß der Präzision ist nicht allein vom Meßgerät abhängig. Die Variabilität im Material, hier die verschiedenen Cigaretten, tragen wesentlich bei zur Streuung der Meßwerte. Die Beschaffenheit der Cigaretten beeinflußt die Präzision der Messungen.

Bezeichnen wir mit R den Referenzwert, so stellt die Differenz ( $\mu$ -R) den sog. systematischen Fehler des Meßvorgangs dar:

Genauigkeit = Präzision + Systematischer Fehler 
$$(x - R)$$
 =  $(x - \mu)$  +  $(\mu - R)$ . [2]

Existiert kein Referenzwert und weichen die Messungen zweier Rauchmaschinen voneinander ab, dann wollen wir von systematischen Abweichungen  $\mu_1 - \mu_2$  sprechen, wobei  $\mu_1$  der Mittelwert der ersten,  $\mu_2$  der Mittelwert der zweiten Maschine darstellen. Zur statistischen Charakterisierung dieser systematischen Abweichungen genügen die empirischen Mittelwerte

$$\bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i \text{ und } \bar{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i;$$
 [3]

dabei sind  $n_1$  und  $n_2$  die Umfänge der Stichproben beider Rauchmaschinen. Ein Schätzwert für die systematischen Abweichungen beträgt  $(\bar{x} - \bar{y})$ .

Nicht unerwähnt bleibe die Tatsache, daß sowohl die Präzision als auch die systematischen Unterschiede in den Mittelwerten nicht unbedingt als feste Größen angesehen werden dürfen:

a) Entscheidend kann die Abhängigkeit vom Ausmaß
des zu messenden Wertes μ sein. Oft nimmt beispielsweise die Standardabweichung proportional
mit dem Ausmaß der Messung zu, so daß die Variationskoeffizienten

$$V = \frac{s}{\bar{x}}$$
 [4]

konstant bleiben und die Präzision oder die systematischen Fehler direkt vom Ausmaß der Messung abhängig werden.

- b) Weiter können die Umweltbedingungen die Genauigkeit der Messungen beeinflussen. Dem ist durch eine umfassende Standardisierung des Meßvorganges zu begegnen. Eventuell ist es auch möglich, nicht standardisierbare Störgrößen zumindest als Kovariablen bei der Auswertung miteinzubeziehen (dies erfordert Kovarianzanalysen an Stelle von Varianzanalysen).
- c) Die Genauigkeit kann von der Art des Materials, hier etwa die Cigarettenmarke, abhängen.

<sup>\*</sup> Eingegangen am 24. September 1971.

## 3. VERGLEICH DER GENAUIGKEIT ZWEIER RAUCHMASCHINEN BEI GLEICHER ANZAHL ABGERAUCHTER CIGARETTEN

Pro Analysengang seien gleichviel Cigaretten mit beiden Rauchmaschinen abgeraucht worden. Der Vergleich der Präzision erfolgt bei Annahme normalverteilter Analysenwerte durch das Prüfen des Quotienten beider Varianzen.

$$F = \frac{s_1^2}{s_2^2}$$
 [5]

ist bei Gültigkeit der Nullhypothese  $\sigma_1 = \sigma_2$ , d. h. bei Übereinstimmung in den theoretischen Varianzen oder Standardabweichungen, F-verteilt mit  $(n_1-1)$  und  $(n_2-1)$  Freiheitsgraden. Die theoretischen Signifikanzgrenzen zu gebräuchlichen Sicherheitsschwellen sind in einschlägigen Tabellenwerken oder Lehrbüchern der Statistik angegeben. Ist F < 1, so ist der Reziprokwert

$$\frac{1}{F} = \frac{s_2^2}{s_1^2}$$
 [6]

ebenfalls F-verteilt, jedoch mit (n<sub>2</sub>-1) und (n<sub>1</sub>-1) Freiheitsgraden. Die erste Zahl bezieht sich auf die Varianz im Zähler, die zweite auf die Varianz im Nenner des Quotienten.

Ist die Präzision zweier Maschinen nicht wesentlich voneinander verschieden, so sind wir zur Beurteilung der systematischen Abweichungen berechtigt, die Annahme  $\sigma_1 = \sigma_2$  zu treffen. Unter dieser Voraussetzung berechnet man die nach "Student" benannte Größe

$$t = \frac{\overline{x} - \overline{y}}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} , \qquad [7]$$

mit

$$s^2 = \frac{1}{n_1 + n_2 - 2} [(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2].$$

Unter Gültigkeit der Nullhypothese  $\mu_1=\mu_2$ , d. h. bei Übereinstimmung in den theoretischen Mittelwerten, können Signifikanzgrenzen zu vorgegebenen Sicherheitsschwellen und in Abhängigkeit der Anzahl Freiheitsgrade — hier  $(n_1+n_2-2)$  — nachgeschlagen werden.

Werden aber bei zwei Maschinen ungleichviel Cigaretten pro Analysengang abgeraucht, so sind die hier dargebrachten Testansätze nicht mehr gültig. Eine Verallgemeinerung des t-Testes ist aber trotzdem noch möglich.

# 4. VERGLEICH DER GENAUIGKEIT ZWEIER RAUCHMASCHINEN BEI UNGLEICHER ANZAHL ABGERAUCHTER CIGARETTEN

Aus je  $k_1$  Cigaretten gewinne man  $n_1$  Analysenmittelwerte  $x_1, x_2, \ldots, x_{n1}$  mit dem theoretischen Mittelwert  $\mu_1$  und der theoretischen Varianz  $\sigma_1{}^2/k_1$  für die erste Rauchmaschine. Aus je  $k_2$  Cigaretten erhalte man weiter  $n_2$  Analysenmittelwerte  $y_1, y_2, \ldots, y_{n2}$  mit dem Mittelwerte  $\mu_2$  und der Varianz  $\sigma_2{}^2/k_2$  für die zweite Rauchmaschine. Ferner nehmen wir an, daß die Meßwerte Stichproben einer Normalverteilung darstellen. Sollen die Varianzen  $\sigma_1{}^2$  und  $\sigma_2{}^2$  miteinander verglichen werden, so können die Analysenwerte so modifiziert werden, daß zwei

neue Zahlenreihen entstehen, die direkt miteinander verglichen werden können. Dazu werden die x-Werte mit  $\sqrt{k_1}$ , die y-Werte mit  $\sqrt{k_2}$  multipliziert. Die neuen Zahlenwerte

$$x_{i}' = \sqrt{k_1} x_i$$
  $(i = 1, 2, ..., n_1)$  [8]

haben einen theoretischen Mittelwert  $\sqrt{k_1 \mu_1}$  und eine Varianz  $\sigma_1^2$ .

$$y_{i}' = \sqrt{k_2} y_i$$
  $(j = 1,2,...,n_2)$  [9]

besitzen den Mittelwert  $\sqrt{k_2} \mu_2$  und die Varianz  $\sigma_2^2$ .

Die Theorie der linearen Modelle, vergleiche dazu etwa Graybill (1), erlaubt den Vergleich verschiedener statistischer Modelle. Unter der Nullhypothese  $\sigma_1 = \sigma_2$ , m. a. W. Übereinstimmung in den Varianzen, ist der Quotient

$$F = \frac{s_x^2}{s_v^2}$$
 [10]

mit

$$s_{x}^{2} = \frac{1}{n_{1}-1} \sum_{i=1}^{n_{1}} (x_{i}' - \overline{x}')^{2}$$

und

$$s_y^2 = \frac{1}{n_2-1} \sum_{j=1}^{n_2} (y_j' - \overline{y}')^2$$

F-verteilt mit  $(n_1-1)$  und  $(n_2-1)$  Freiheitsgraden. Zudem gilt  $s_x$ ,  $s_y$  =  $s_y$  und  $s_y$ ,  $s_y$  =  $s_y$ , so daß sich F direkt aus den Maßzahlen der ursprünglichen Analysenwerte bestimmen läßt zu

$$F = \frac{k_1}{k_2} \cdot \frac{s_x^2}{s_v^2} .$$
 [11]

Ist  $F \leq 1$ , so ist der Reziprokwert 1/F wiederum F-verteilt, jetzt mit  $(n_2-1)$  und  $(n_1-1)$  Freiheitsgraden. Sind die Standardabweichungen  $s_1$  und  $s_2$  nicht wesentlich voneinander verschieden, so dürfen wir die Mittelwerte unter der Annahme  $\sigma_1 = \sigma_2$  miteinander vergleichen. Unter dieser Voraussetzung können die Mittelwerte unter einer Verallgemeinerung des t-Testes verglichen werden. Ist nämlich  $\mu_1 \neq \mu_2$ , so sind  $\overline{x}$  resp.  $\overline{y}$  geeignete Schätzwerte für  $\mu_1$  und  $\mu_2$ . Die Summe der Quadrate um dieses Alternativmodell ist

$$S_{A} = \sum_{i=1}^{n_{1}} (x_{i}' - \overline{x}')^{2} + \sum_{i=1}^{n_{2}} (y_{i}' - \overline{y}')^{2}; \quad [12]$$

oder in den ursprünglichen Größen ausgedrückt

$$S_{A} = k_{1} \sum_{i=1}^{n_{1}} (x_{i} - \overline{x})^{2} + k_{2} \sum_{j=1}^{n_{2}} (y_{j} - \overline{y})^{2}. [13]$$

Ist dagegen  $\mu_1 = \mu_2 = \mu$ , so kann ein Schätzwert für  $\mu$  nach der Maximum-Likelihood-Methode oder nach der Methode der kleinsten Quadrate erhalten werden. Die Bestimmungsgleichung hierzu ist:

$$\sum_{i=1}^{n_1} (x_i' - \sqrt{k_1} \mu)^2 + \sum_{i=1}^{n_2} (y_i' - \sqrt{k_2} \mu)^2 = Min.$$

[14]

$$\hat{\mu} = \frac{k_1 n_1 \overline{x} + k_2 n_2 \overline{y}}{k_1 n_1 + k_2 n_2}$$
 [15]

und als Summe der Quadrate um dieses Nullmodell

$$S_0 = \sum_{i=1}^{n_1} (x_i' - \sqrt{k_1 \mu})^2 + \sum_{j=1}^{n_2} (y_j' - \sqrt{k_2 \mu})^2. \quad [16]$$

So ist nun stets größer als SA. Je größer die Differenz

$$D = S_0 - S_A$$
, [17]

desto größer sind auch die Unterschiede in den Mittelwerten zweier Maschinen. Aus den Größen  $S_A$ ,  $S_0$  und D bilden wir eine Varianzanalyse gemäß Tabelle 1.

Tabelle 1. Varianzanalyse.

Varianz- komponente	Summe der Quadrate	Freiheits- grade	Durchschnitts- quadrat
Nullmodell $\mu_1 = \mu_2$	s <sub>o</sub>	n <sub>1</sub> + n <sub>2</sub> - 1	. •
Alternativmodel $\mu_1 \neq \mu_2$	S <sub>A</sub>	$n_1 + n_2 - 2$	$S_A/(n_1+n_2-2)$
Systematische Abweichungen			

 $d = S_0 - S_A$ 

Der Quotient

Mittelwerten

$$F = \frac{D}{S_A} (n_1 + n_2 - 2)$$
 [18]

ist unter Gültigkeit der Nullhypothese F-verteilt mit einem und  $(n_1+n_2-2)$  Freiheitsgraden. Für die numerische Berechnung der Summen der Quadrate ermittelt man mit Vorteil  $S_A$  und D, da sich D leicht vereinfachen läßt zu

$$D = \frac{k_1 k_2 n_1 n_2 (\bar{x} - \bar{y})^2}{k_1 n_1 + k_2 n_2} .$$
 [19]

Da das Quadrat der t-Testgröße mit  $(n_1+n_2-2)$  Freiheitsgraden eine F-Verteilung mit einem und  $(n_1+n_2-2)$  Freiheitsgraden befolgt, können wir an Stelle der Varianzanalyse einen verallgemeinerten t-Test zum Prüfen der Unterschiede in den Mittelwerten aufstellen.

$$t = \frac{\bar{x} - \bar{y}}{s} \sqrt{\frac{k_1 k_2 n_1 n_2}{k_1 n_1 + k_2 n_2}}$$
, [20]

mit

$$s^{2} = \frac{1}{n_{1} + n_{2} - 2} \left[ k_{1} \sum_{i=1}^{n_{1}} (x_{i} - \bar{x})^{2} + k_{2} \sum_{i=1}^{n_{2}} (y_{i} - \bar{y})^{2} \right],$$

ist t-verteilt mit ( $n_1 + n_2$ -2) Freiheitsgraden. Ist  $k_2/k_1 = m$  etwa ganzzahlig, so kann [20] günstig umgeformt werden zu

$$t = \frac{\overline{x} - \overline{y}}{s} \sqrt{\frac{m n_1 n_2}{n_1 + m n_2}}$$
 [21]

mit

$$s^2 = \frac{1}{n_1 + n_2 - 2} \left[ \sum_{i=1}^{n_1} (x_i - \bar{x})^2 + m \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \right].$$

Ist ein errechneter t-Wert größer als die Signifikanzgrenze zu einer vorgegebenen Sicherheitsschwelle, so betrachten wir die systematische Abweichung als statistisch signifikant, d. h. daß die Nullhypothese  $\mu_1 = \mu_2$  abgelehnt wird.

#### 5. EIN BEISPIEL

Vergleich der Rauchmaschinen P. M. 20 Port und RM 20/68

#### 5.1 Einleitung

Es wurden  $k_1$  bzw.  $k_2$  Cigaretten in einer Rauchfalle abgeraucht:

P. M. 20 Port 5 Cigaretten 
$$(k_1 = 5)$$
  
RM 20/68 20 Cigaretten  $(k_2 = 20)$ 

Es wurden bestimmt:

Feuchtes Rauchkondensat (TPM) Trockenes Rauchkondensat (DPM)

und

Rauchnikotin (SN)\*\*

für die 3 Cigarettenarten:

Strangcigarette (N) Filtercigarette M (M)

Filtercigarette F (F)

Im folgenden werden Präzision und systematische Abweichungen beider Rauchmaschinen, bezogen auf einen Cigarettenmittelwert, miteinander verglichen.

Für die Beurteilung von Testgrößen wird immer eine Sicherheitsschwelle von 5%0 angesetzt. Wir kennzeichnen eine signifikante Abweichung ( $P \le 0.05$ ) jeweils mit \*. Eine nicht signifikante Abweichung (P > 0.05) mit (–).

#### 5.2 Vergleich der Präzision

5.2.1 Vergleich der totalen Variabilität: In einer ersten Phase wurden die Varianzen für die 3 Analysenwerte TPM, DPM und SN und für jede Cigarettenart N, M und F auf Unterschiede in der Präzision hin geprüft.

Die Ergebnisse – in Tabellen 2 und 3 zusammengestellt – zeigen, daß für TPM nur für die M-Cigaretten signi-

Tabelle 2. Varianzen nach Cigarettenart und Maschinentyp (nichtgewichtet).

	P.N	P.M. 20 Port		RM 20/68		
	N	M.	F	N	М	F
TPM	1,112	0,936	0,936	0,1466	0,0800	0,2466
DPM SN	0,905 0,00237	0,833 0,00324	0,702 0,00208	0,1088 0,00937		•

<sup>\*\*</sup> Die analytisch gefundenen Nikotinwerte sind durch einen systematischen Analysenfehler verfälscht. Sie werden ausschließlich zur Illustration angeführt.

Tabelle 3. F-Testgröße des Vergleichs der Varianzen der beiden Maschinentypen +.

	N	M	F	
ТРМ	1,901 (—)	2,906 *	1,901 (—)	
DPM	2,115 ()	1,075 ()	0,904 ()	
SN	0,063 *	0,212 *	0,114 *	

<sup>+ \*</sup> bedeutet statistisch signifikant zur Sicherheitsschwelle von 5 %.
(-) bedeutet statistisch nicht signifikant zur Sicherheitsschwelle von 5 %.

fikante Unterschiede in der Präzision auftreten, daß dagegen für DPM alle Unterschiede statistisch unbedeutend sind. Für SN beobachtet man für alle Cigarettenarten signifikante Unterschiede zugunsten einer kleineren Varianz für die Rauchmaschine P. M. 20 Port.

## 5.3 Vergleich der Mittelwerte beider Maschinentypen

Für die wichtigsten Größen TPM und DPM werden nach Formel [21] mit m = 4 die Mittelwerte — die sog. systematischen Abweichungen messend — miteinander verglichen.

Tabelle 4 zeigt, daß diese Mittelwerte bis auf einen einzigen durchweg signifikant aussielen. Für M und F liefert die P. M. 20 Port etwas größere Werte als die RM 20/68; dagegen sind die Mittelwerte der N-Cigarette für die RM 20/68 etwas größer ausgefallen.

Tabelle 4. Mittelwerte und t-Testgröße zum Vergleich der beiden Maschinentypen auf systematische Abweichungen.

	P.M. 20 Port	RM 20/68	t-Testgröße und Wertung
TPM N	26,13	26,38	— 1,27 ( <b>–</b> )
M	25,44	24,46	+ 4,81 *
F	20,75	19,87	+ 4,81 *
DPM N	23,70	24,50	4,29 *
M	22,38	21,63	+ 3,69 *
F	18,91	18,29	+ 3,65 *

#### ZUSAMMENFASSUNG

Bei verschiedenen Rauchmaschinen werden oft ungleich viel Cigaretten in einer Rauchfalle abgeraucht. Dies hat zur Folge, daß der ungleichen Zahl von Cigaretten wegen auch die Variabilitäten verschieden sind. Für den Vergleich von Mittelwerten darf aus diesem Grunde der bekannte t-Test nicht direkt angewendet werden. Es wird gezeigt, daß ein verallgemeinerter t-Test unter

modifizierten Voraussetzungen trotzdem meistens noch durchführbar ist; dann nämlich, wenn die Standardabweichungen pro Cigarette nicht wesentlich verschieden sind.

#### SUMMARY

The performances of smoking machines of different types cannot be compared directly if different numbers of cigarettes are smoked into one trap, because the respective variabilities are depending on the number of cigarettes smoked into one trap. For the comparison of means, the t-test can therefore not be used directly. It is demonstrated that a generalized t-test can not-withstanding be applied to this problem in most cases, in particular when the standard deviations per cigarette are of the same order of magnitude.

#### **RESUME**

Les performances de machines à fumer de types différents ne se laissent pas comparer sans autre si le nombre de cigarettes fumées dans un piège est différent, car ceci a une influence directe sur les variabilités respectives. En vue d'une comparaison des résultats moyens obtenus sur les deux machines le test de t ne peut donc pas être employé sans autre. Il est démontré ici que l'application d'un test de t généralisé est néanmoins possible dans la plupart des cas, en particulier lorsque les déviations standard par cigarette ne diffèrent que légèrement.

#### LITERATUR

- Graybill, F. A.: An introduction to linear statistical models; McGraw-Hill Book Company, New York, 1961.
- 2. Lipp, G.: Vergleich der Rauchausbeute von drei verschiedenen Rauchmaschinen (I); Beitr. Tabakforsch. 5 (1969) 39–42.
- 3. Wiley, R. M., und Ferri, J. A.: Beitr. Tabakforsch. 4 (1967) 59-64.

#### Anschrift des Verfassers:

Universität Bern, Institut für Mathematische Statistik und Versicherungslehre, Sidlerstraße 5, CH-3000 Bern, Schweiz.