proceedings of conference on foundational questions in statistical inference

aarhus,
may 7-12, 1973

editors:

ole barndorff-nielsen
preben blæsild
geert schou

ON SPECIFIC DISTRIBUTIONS FOR TESTING OF HYPOTHESES

G. Rasch
University of Copenhagen

In this contribution some examples are mentioned of probabilistic statements that have the property of being specific for, i.e. necessary and sufficient conditions for the validity of, the statistical hypotheses from which they are deduced. It is contended that, whenever possible, statements of this kind should be chosen as basis for tests of statistical hypotheses.

1. The Multiplicative Poisson Model (MPM)

Although it has been used in various connections I may present the MPM in the same kind of context as Martin-Löf dealt with, accidents with person injuries, in this case having occurred on Danish roads during some months of the years 1961/64. The roads were grouped into 24 categories according to road profile, to their passing through areas with dense or dispersed building-up and to sections with and without crossings or junctions. The daily observations were grouped according to the days of the week while Easter and Whit-Sunday, etc. were excluded.

Within each framework thus established the data may be presented in a rectangular matrix:

(1.1) roadcategories

$r \backslash t$	days $1,\ldots,t,\ldots,T$	total
1	.	
.	.	
.	.	
.	.	
r	a_{rt}	a_{r+}
.	.	
.	.	
.	.	
R	.	
total	a_{+t}	a_{++}

The model mentioned assumed that

$$(1.2) \qquad p\{a_{rt}\} = e^{-\lambda_{rt}} \cdot \frac{\lambda_{rt}^{a_{rt}}}{a_{rt}!}, \text{ where } \lambda_{rt} = \rho_r\, \theta_t,$$

ρ_r characterizing the danger of driving on the road category r, θ_t that of driving during the day t.

Assuming further stochastic independence for fixed set of parameters, we get for the whole set of observations indicated

$$p\{((a_{rt}))\} = e^{-\rho_+\,\theta_+} \cdot \frac{\prod\limits_{(r)} \rho_r^{a_{r+}} \cdot \prod\limits_{(t)} \theta_t^{a_{+t}}}{\prod\limits_{(r)} \prod\limits_{(t)} a_{rt}!},$$

from which the following distributions are easily derived

$$p\{a_{++}\} = e^{-\rho_+\,\theta_+} \frac{(\rho_+\,\theta_+)^{a_{++}}}{a_{++}!},$$

$$p\{(a_{r+}),(a_{+t}),a_{++}\} = p\{(a_{r+})|a_{++}\}\; p\,\{(a_{+t})|a_{++}\}\; p\,\{a_{++}\},$$

$$(1.3) \qquad p\{(a_{r+})|a_{++}\} = \frac{a_{++}!}{\prod\limits_{(r)} a_{r+}!} \cdot \prod_{(r)} \left(\frac{\rho_r}{\rho_+}\right)^{a_{r+}},$$

$$(1.4) \qquad p\{(a_{+t})|a_{++}\} = \frac{a_{++}!}{\prod\limits_{(t)} a_{+t}!} \cdot \prod_{(t)} \left(\frac{\theta_t}{\theta_+}\right)^{a_{+t}},$$

$$(1.5) \qquad p\{((a_{rt}))|(a_{r+}),(a_{+t}),\ a_{++}\} = \frac{\prod\limits_{(r)} a_{r+}!\ \prod\limits_{(t)} a_{+t}!}{a_{++}!\ \prod\limits_{(r)}\prod\limits_{(t)} a_{rt}!} .$$

Accordingly this model allows for separating the evaluation of the road parameters through the conditional distribution (1.3) from that of the day parameters through (1.4) and from a parameterfree model control through (1.5).

In fact, (1.5) being a consequence of the model, its validity is a <u>necessary condition</u> for the model to hold; thus consequences of it may be used in attempts at exploring the validity of the model. But such procedures may raise a critical question: How specific would such tests be ?

Of course, tests as such can never <u>prove</u> any hypothesis, but the probability statement behind a test may either be just some consequence of the hypothesis in question, i.e. expressing a condition that is <u>necessary</u> for the hypothesis to be true, or it may be a statement from the correctness of which the truth of the hypothesis would follow, i.e. a condition that is <u>sufficient</u> to imply the hypothesis.

<u>I suggest that, whenever possible, the background statement should express a condition that is both necessary and sufficient for the validity of the hypothesis in question</u>.

It would seem possible, though somewhat intricate, to prove that (1.5) - or even a simple consequence of it - has the property of being specific for truth of the basic model (1.2); I shall, however, confine myself to another condition which is also specific, but easier to handle, namely an inversion to a special case of the statement (1.3), pertaining to just one column at a time in the matrix (1.1):

$$(1.6) \qquad p\{a_{1t},\ldots,a_{Rt} \mid a_{+t} = c_t\} = \binom{c_t}{a_{1t},\ldots,a_{Rt}} \cdot \prod_{r=1}^{R} \left(\frac{\rho_r}{\rho_+}\right)^{a_{rt}} .$$

This is a multinomial distribution, which, however, possesses the same parameters for all c_t's.

Theorem 1.* Assume that all the variates a_{rt}, $r = 1,\ldots,R$, $t = 1,\ldots,T$ for arbitrary R and T are stochastically independent. Assume furthermore that the distributions of the vectors $a_{*t} = (a_{1t},\ldots,a_{Rt})$ have the properties in common with the multinomial distributions (1.6), that the mean values for any c_t are proportional to c_t, i.e.

$$\mathfrak{m}\{a_{rt} \mid c_t\} = \alpha_r\, c_t (0 \le \alpha_r \le 1,\ \alpha_+ = 1)$$

and that the same holds for the variances and covariances:

$$\mathrm{V}\{a_{rt} \mid c_t\} = \alpha_r(1-\alpha_r)\, c_t\ ,$$

$$\mathrm{V}\{a_{rt},\ a_{st} \mid c_t\} = -\alpha_r\alpha_s \cdot c_t,\quad r \neq s.$$

Then the set of variates a_{rt}, $r = 1,\ldots,R$, $t = 1,\ldots,T$ must follow the Multiplicative Poisson Model with the road parameters α_r.

The idea in the proof may be sufficiently well demonstrated by the case R = 2. **

---

*This result is a refinement of a theorem due to Bol'shev (1965). See also the next footnote.

**In this case, the result follows immediately from a general, closely related theorem due to Bolger and Harkness (1965). Their method of proof is quite similar to the one given here.

For probability generating functions, I find the following notations convenient. For any enumerative variate a with probabilities p{a} I write

$$\pi\{{}^{a}_{x}\} = \sum_{(a)} p\{a\}\ x^{a}$$

and for a pair (a,b)

$$\pi\{{}^{a,b}_{x,y}\} = \sum_{(a,b)} p\{a,b\}\ x^{a}\ y^{b}.$$

From p{a,b} = p{b|a}p{a} it follows that

$$\pi\{{}^{a,b}_{x,y}\} = \sum_{(a)} \Big(\sum_{(b)} p\{b|a\}\ y^{b}\Big)\ p\{a\}\ x^{a}$$

$$= \sum_{(a)} \pi\{{}^{b}_{y}|a\}\ p\{a\}\ x^{a}.$$

In particular for a + b = c

$$\pi\{{}^{a}_{xz},\ {}^{b}_{yz}\} = \sum_{(c)} \pi\{{}^{a,b}_{x,y}|c\}\ p\{c\}\ z^{c}.$$

If furthermore a and b are stochastically independent this becomes

$$\pi\{{}^{a}_{xz}\}\ \pi\{{}^{b}_{yz}\} = \sum_{(c)} \pi\{{}^{a,b}_{x,y}|c\}\ p\{c\}\ z^{c}\ ,$$

from which, on differentiation once and twice, afterwards putting x = y = 1, the following results are obtained

$$(1.7) \qquad z\ \pi'\{{}^{a}_{z}\}\ \pi\{{}^{b}_{z}\} = \sum_{(c)} \mathfrak{m}\{a|c\}\ p\{c\}\ z^{c}$$

and

$$(1.8) \qquad z^{2}\ \pi'\{{}^{a}_{z}\}\ \pi'\{{}^{b}_{z}\} = \sum_{(c)} \mathfrak{m}\{ab|c\}\ p\{c\}\ z^{c}\ ,$$

$\pi'\{{}^{a}_{z}\}$ denoting $\dfrac{\partial \pi\{{}^{a}_{z}\}}{\partial z}$.

Now insert the condition

$$\mathfrak{m}\{a|c\} = \alpha\ c, \quad c = 0,1,2,\ldots.$$

into (1.7). Then we get

$$\pi'\{\genfrac{}{}{0pt}{}{a}{z}\}\ \pi\{\genfrac{}{}{0pt}{}{b}{z}\} = \alpha \sum_{(c)} c\ p\{c\}\ z^{c-1} = \alpha\ \pi'\{\genfrac{}{}{0pt}{}{c}{z}\}\ ,$$

and since

$$(1.9) \qquad \pi\{\genfrac{}{}{0pt}{}{c}{z}\} = \pi\{\genfrac{}{}{0pt}{}{a,b}{z,z}\} = \pi\{\genfrac{}{}{0pt}{}{a}{z}\}\ \pi\{\genfrac{}{}{0pt}{}{b}{z}\}$$

this implies

$$(1.10.a) \qquad \frac{\pi'\{\genfrac{}{}{0pt}{}{a}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{a}{z}\}} = \alpha\ \frac{\pi'\{\genfrac{}{}{0pt}{}{c}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{c}{z}\}}$$

and by analogy

$$(1.10.b) \qquad \frac{\pi'\{\genfrac{}{}{0pt}{}{b}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{b}{z}\}} = (1-\alpha)\ \frac{\pi'\{\genfrac{}{}{0pt}{}{c}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{c}{z}\}}\ .$$

As regard (1.8) the relations

$$V\{a|c\} = -V\{a,b|c\}$$

and

$$V\{a|c\} = \alpha(1-\alpha)\ c$$

imply

$$\mathfrak{m}\{ab|c\} = \alpha(1-\alpha)\ c(c-1).$$

Accordingly (1.8) is equivalent to

$$\pi'\{\genfrac{}{}{0pt}{}{a}{z}\}\ \pi'\{\genfrac{}{}{0pt}{}{b}{z}\} = \alpha(1-\alpha)\ \pi''\{\genfrac{}{}{0pt}{}{c}{z}\}$$

which on division by $\pi\{\genfrac{}{}{0pt}{}{c}{z}\}$ (cf.(1.9)) and insertion of (1.10.a,b) yields

$$\alpha(1-\alpha)\left(\frac{\pi'\{\genfrac{}{}{0pt}{}{c}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{c}{z}\}}\right)^2 = \alpha(1-\alpha)\ \frac{\pi''\{\genfrac{}{}{0pt}{}{c}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{c}{z}\}}$$

or

$$\frac{\pi'\{\genfrac{}{}{0pt}{}{c}{z}\}}{\pi\{\genfrac{}{}{0pt}{}{c}{z}\}} = \frac{\pi''\{\genfrac{}{}{0pt}{}{c}{z}\}}{\pi'\{\genfrac{}{}{0pt}{}{c}{z}\}}\ ,$$

i.e.

$$\log \pi\{{}^{c}_{z}\} = \log \pi'\{{}^{c}_{z}\} - \log \mathfrak{m}\{c\} ,$$

from which

$$\pi\{{}^{c}_{z}\} = e^{\delta(z-1)} , \text{ with } \delta = \mathfrak{m}\{c\},$$

follows. Then, according to (1.10.a,b)

$$\pi\{{}^{a}_{z}\} = e^{\alpha\delta(z-1)} , \quad \pi\{{}^{b}_{z}\} = e^{(1-\alpha)\delta(z-1)}$$

Referring to the situation described in the theorem (for $R = 2$) we may write

$$\pi\{{}^{a_1 t}_{z}\} = e^{\alpha_1\delta_t(z-1)} , \quad \pi\{{}^{a_2 t}_{z}\} = e^{\alpha_2\delta_t(z-1)}$$

since

$$\delta_t = \mathfrak{m}\{a_{1t}\} + \mathfrak{m}\{a_{2t}\}$$

varies with t. This completes the proof for $R = 2$.

As a consequence of Theorem 1 we may now turn the argument the other way round:

Corollary: If the proportionality conditions are fulfilled, but notwithstanding the distributions $p\{a_{1t},\ldots,a_{Rt}|c_t\}$ are not multinomial, then the a_{rt}'s cannot be stochastically independent.

2. Fisher's "exact test"

Fisher's classical statement on independent, binomially distributed variates says that if $\theta_1 = \theta_2$ in the distributions

$$p\{a_1|n_1\} = \binom{n_1}{a_1} \theta_1^{a_1} (1-\theta_1)^{n_1 - a_1}, \quad p\{a_2|n_2\} = \binom{n_2}{a_2} \theta_2^{a_2} (1-\theta_2)^{n_2-a_2},$$

then the following parameterfree conditional distribution holds:

$$(2.1) \qquad p\{a_1,a_2|a_+,n_1,n_2\} = \frac{\binom{n_1}{a_1}\binom{n_2}{a_2}}{\binom{n_1 + n_2}{a_1 + a_2}} .$$

Occasionally arguments on the adequacy of such conditional testing has been going on, but in this particular case I think the answer has been given by the demonstration by Patil and Seshadri (1964) and Menon (1966) that the condition (2.1) is not only necessary, but also sufficient for $\theta_1 = \theta_2$.

By the same type of argument as used in Section 1 we may even prove a corresponding kind of inversion to (2.1):

Theorem 2. If the variates a_1 for given n_1 and a_2 for given n_2 are stochastically independent and if for any value of

$$c = a_1 + a_2$$

we have

$$\mathfrak{m}\{a_1 | c_1\ n_1,\ n_2\} = \frac{n_1}{n_+}$$

and

$$\mathfrak{m}\{a_1\ a_2 | c_1\ n_1,\ n_2\} = n_1\ n_2\ \frac{a_+(a_+-1)}{n_+(n_+-1)}, \tag{2.2}$$

then such a parameter θ must exist that

$$p\{a_j | n_j\} = \binom{n_j}{a_j} \theta^{a_j} (1-\theta)^{n_j - a_j}, \qquad j = 1,2 .$$

The mean value part of the proof is practically unchanged, while the second part, in particular due to the factor $\frac{a_+-1}{n_+-1}$ in the variance (2.2), just changes in such a way that we instead of the exponential function referring to the Poisson distribution obtain the binomials characteristic of the generating functions for binomial distributions.

3. Further, related results

In the additional paper, read before the Danish Society for Theoretical Statistics, September 14th, 1971, similar results are obtained for other distributions related to the exponential family.

While all the above-mentioned results refer to conditional distributions, alike propositions holding for marginal distributions are due to Prohorov (1966) and Csörgö and Seshadri (1970).

Acknowledgements

I owe the references given to Barndorff-Nielsen and Martin-Löf.

References

Bolger, E. M. and Harkness, W. L. (1965): Characterizations of some distributions by conditional moments. Ann. Math. Statist. 36, 703-705.

Bol'shev, L. N. (1965): On a characterization of the Poisson distribution and its statistical applications. Theory Probability Appl. 10, 446-456.

Csörgö, M. and Seshadri, V. (1970): On the problem of replacing composite hypotheses by equivalent simple ones. Rev. Int. Statist. Inst. 38, 351-368.

Menon, M. V. (1966): Characterization theorems for some univariate probability distributions. J. Roy. Statist. Soc. B. 28, 143-145.

Patil, G. P. and Seshadri, V. (1964): Characterization theorems for some univariate probability distributions. J. Roy. Statist. Soc. B. 26, 286-292.

Prohorov, Yu. V. (1966): Some characterization problems in statistics. Proc. Fifth Berkeley Symp. Math. Statist. Prob., Vol. I, 341-349.