Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

A CONCEPT OF δ-STOCHASTIC ORDERING AND ITS APPLICATION TO THE COMPETING RISKS MODEL

S. T. RACHEV, A. YU. YAKOVLEV, N. O. KADYROVA

We call a random variable Y δ-stochastically smaller than a random variable X if for a given $\delta \geqq 0$ the following relation between distribution functions holds: $F_X(t)-\delta \leqq F_Y(t)$, $t \in R^1$.

Using this notion within competing risks framework, it is possible to introduce new probabilistic characteristics of each type of failure contribution to the joint effect of two damaging agents.

Confidence bounds for such characteristics are given.

1. Introduction. To construct the confidence bounds for the values of crude survival probabilities $Q_1(t)$ and $Q_2(t)$ at $t=0$ (i. e. $\Pi_1=\mathrm{pr}(T_1<T_2)$ and $\Pi_2=\mathrm{pr}(T_1>T_2)$, T_1 and T_2 are the latent failure times) in our previous communication (S. Rachev, A. Yakovlev, [3]), we proceeded from the inequalities

(1) $$\sigma(F_2, F) \leqq \Pi_1 \leqq 1-\sigma(F_1, F), \quad \sigma(F_1, F) \leqq \Pi_2 \leqq 1-\sigma(F_2, F),$$

where $F_1(t)$ and $F_2(t)$ are the marginal distribution functions for random variables T_1 and T_2, $F(t)$ is the distribution function for the variable $T=\min(T_1, T_2)$, and $\sigma(F', F'')$ is the distance in total variation between two random variables X' and X'' with distributions $F'(x)$ and $F''(x)$ respectively.

Using the inequality $\sigma(F', F'') \geqq \rho(F', F'')$, where $\rho(F', F'')$ is the uniform metric, we have obtained from (1) the following bounds

(2) $$\rho(F_2, F) \leqq \Pi_1 \leqq 1-\rho(F_1, F), \ \rho(F_1, F) \leqq \Pi_2 \leqq 1-\rho(F_2, F),$$

and used them for the derivation of confidence bounds in terms of the empiric distributions $F_1^{(n)}(t)$, $F_2^{(k)}(t)$ and $F^{(m)}(t)$. One must expect that the transition from (1) to (2) results in a highly rough estimation because the metric $\sigma(F', F'')$ is topologically stronger than the metric $\rho(F', F'')$. In other words it is impossible to find continuous strictly increasing on $[0, \infty)$ function φ such that $\varphi(0)=0$ and

$$\rho(F', F'') \geqq \varphi\{\sigma(F', F'')\}.$$

Thus, to improve the estimation procedure it is necessary to proceed from some other characteristics of each type of failure contribution to the effect of combined injury.

Note that

$$\Pi_1=\mathrm{pr}(T_2>T_1)=\mathrm{pr}\{\bigcup_{x \geqq 0}(T_2>x \geqq T_1)\}.$$

This representation of Π_1 suggests to introduce the failure (type I) contribution in the form of some appropriate functional of pr $(T_2>x \geqq T_1)$. Here we shall try the following ones

SERDICA Bulgaricae mathematicae publications. Vol. 15, 1989, p. 28–31.

$$\Pi_1^l := \inf_{x>0} \frac{\text{pr}\,(T_2 > x \geq T_1)}{\text{pr}\,(T_1 > x \geq T_2)} - 1,$$

(3)

$$\Pi_1^u := \sup_{x>0} \frac{\text{pr}\,(T_2 > x \geq T_1)}{\text{pr}\,(T_1 > x \geq T_2)} - 1.$$

The indicator Π_1^l seems to be contensive because of its relation to the certain kind of two random variables ordering.

2. The notion of δ-stochastic ordering of two random variables. We shall call a random variable Y δ-stochastically smaller than a variable X and write $X \underset{\delta\text{-st}}{>} Y$, if for a given $\delta \geq 0$ and all $t \in R^1$

$$F_X(t) - \delta \leq E_Y(t),$$

where $F_X(t)$ and $F_Y(t)$ are the distribution functions for the variables X and Y, respectively.

At $\delta=0$ this notion coincides with the well-known (see: § 1.2, D. Stoyan [4]) definition of the ordering relation $\underset{\leq}{(1)}$ for two random variables. The usefulness of the generalization $\underset{\leq}{(1)}$ for testing the stochastic inequalities can be illustrated by the following examples.

Let $F_{X_n}(t)$ be a sequence of distribution functions for random variables X_n and, $F_{Y_n}(t)$ be a simillar sequence for random variables Y_n. From the uniform convergence i. e. the convergence in metric $\rho(F', F'')$, of the sequences $F_{X_n}(t)$ and $F_{Y_n}(t)$ to the distributions $F_X(t)$ and $F_Y(t)$, where $F_X(t) \leq F_Y(t)$, it does not follow that there exists such n_0 that the relation $X_n \underset{\geq}{(1)} Y_n$ (or $F_{X_n} \leq_F F_{Y_n}$) is valid for all values $n \geq n_0$.

However, in this case it can be stated that $X_n \underset{\delta_n\text{—st}}{>} Y_n$, if δ_n is chosen in the following manner:

$$\delta_n = \delta + p(F_{X_n}, F_X) + \rho(F_{Y_n}, F_Y).$$

It is not difficult to obtain the statistical test for the hypothesis: $X \underset{\delta\text{—st}}{>} Y$. Let $\Phi^{(n)}(x)$ and $R^{(k)}(x)$ be the empirical distribution functions corresponding to $\Phi(x) = \text{pr}\,(X \leq x)$, $R(x) = \text{pr}\,(Y \leq x)$, and assume that

$$\Phi^{(n)}(x) - \delta \leq R^{(k)}(x)$$

for some positive $\delta = \delta(n, k)$ and all values of x.

Then we may write

$$\Phi(x) - R(x) \leq \delta + \rho(\Phi^{(n)}, \Phi) + \rho(R^{(k)}, R),$$

and consequently for any $c_1 > 0$ and $c_2 > 0$

$$\text{pr}\,\{X \underset{\delta + c_1 + c_2 - \text{st}}{>} Y\} \geq \text{pr}\,\{\rho(\Phi^{(n)}, \Phi) \leq c_1,\ \rho(R^{(k)}, R) \leq c_2\}.$$

Using the bounds for joint distribution at fixed marginal distributions (W. Hoeffding, [2]; M. Frechet, [1]), we finally obtain

$$\text{pr}\,(X \underset{\delta + c_1 + c_2 - \text{st}}{>} Y) \geq \max\{G^{(n)}(c_1) + G^{(k)}(c_2) - 1, 0\},$$

where $G^{(n)}(x)$ and $G^{(k)}(x)$ are the Kolmogorov distributions for samples of the sizes n and k respectively. Hence, we have derived the Kolmogorov test for the statistical hypothesis: $X \underset{\delta\text{—st}}{>} Y$.

3. Application to the competing risks model. Returning to the competing risks it is easy to prove the following statement.

In order that the random latent time T_2 be δ-stochastically smaller than the latent time T_1 it is sufficient that for any $\varepsilon \geqq 0$ the following condition is satisfied

$$\Pi_2^l = \inf_{x>0} \frac{\operatorname{pr}(T_1 > x \geqq T_2)}{\operatorname{pr}(T_2 > x \geqq T_1)} - 1 \geqq -\varepsilon. \tag{4}$$

Let us rewrite the expression for Π_2^l as follows

$$\Pi_2^l = \inf_{x>0} \frac{\operatorname{pr}(T_1 > x) - \operatorname{pr}(T_1 > x,\ T_2 > x)}{\operatorname{pr}(T_2 > x) - \operatorname{pr}(T_1 > x,\ T_2 > x)} - 1 = \inf_{x>0} \frac{F_2(x) - F_1(x)}{F(x) - F_2(x)}.$$

It follows from the last expression and the inequality (4) that if $\varepsilon \geqq 0$ then for all $x > 0$

$$F_2(x) - F_1(x) \geqq -\varepsilon \sup_{x>0} [F(x) - F_2(x)],$$

and consequently $F_2(x) \geqq F_1(x) - \delta$, where $\delta = \varepsilon \sup_{x>0} [F(x) - F_2(x)] \geqq 0$. For $\varepsilon < 0$ we have $F_2(x) - F_1(x) \geqq -\varepsilon \inf_{x>0} [F(x) - F_2(x)] = 0$, and $F_2(x) \geqq F_1(x)$, i. e. $\delta = 0$.

Hence there is the additional reason for the value of Π_2^l using: it serves as an indicator of the relation $\overset{>}{\delta\text{-}st}$ between the random variables T_1 and T_2. It is not difficult to find the lower confidence bound for this indicator.

Let $F_1^{(n)}(t)$, $F_2^{(k)}(t)$, $F^{(m)}(t)$ be empirical distributions constructed by use of complete samples of the sizes n, k, m respectively. Then we have the inequality

$$\Pi_2^l \geqq \inf_{t>0} \frac{F_2^{(k)}(t) - \rho(F_2^{(k)},\ F_2) - F_1^{(n)}(t) - \rho(F_1^{(n)},\ F_1)}{F^{(m)}(t) + \rho(F^{(m)},\ F) - F_2^{(k)}(t) + p(F_2^{(k)},\ F_2)}.$$

Now the lower confidence bound for Π_2^l is readily obtained

$$\operatorname{pr}\left\{\Pi_2^l \geqq \inf_{t>0} \frac{F_2^{(k)}(t) - F_1^{(n)}(t) - c_2 - c_1}{F^{(m)}(t) - F_2^{(k)}(t) + c_2 + c}\right\} \geqq G(c_1,\ c_2,\ c),$$

where

$$G(c_1,\ c_2,\ c) = \max\{G^{(k)}(c_2) + G^{(n)}(c_1) + G^{(m)}(c) - 2,\ 0\}.$$

Similarly we introduce another indicator

$$\Pi_2^u = \sup_{x>0} \frac{\operatorname{pr}(T_1 > x \geqq T_2)}{\operatorname{pr}(T_2 > x \geqq T_1)} - 1.$$

It is natural to look for the upper confidence bound for Π_2^u. We get it as follows

$$\operatorname{pr}\left\{\Pi_2^u \leqq \sup_{t>0} \frac{F_2^{(k)}(t) - F_1^{(n)}(t) + c_2 + c_1}{F^{(m)}(t) - F_2^{(k)}(t) - c_2 - c}\right\} \geqq G(c_1,\ c_2,\ c).$$

For the characteristics Π_1^l and Π_1^u defined in (3) the bounds have the form

$$\operatorname{pr}\left\{\Pi_1^l \geqq \inf_{t>0} \frac{F_1^{(n)}(t) - F_2^{(k)}(t) - c_1 - c_2}{F^{(m)}(t) - F_1^{(n)}(t) + c_1 + c}\right\} \geqq G(c_1,\ c_2,\ c)$$

$$\operatorname{pr}\left\{\Pi_1^u \leqq \sup_{t>0} \frac{F_1^{(n)}(t) - F_2^{(k)}(t) + c_1 + c_2}{F^{(m)}(t) - F_1^{(n)}(t) - c_1 - c}\right\} \geqq G(c_1,\ c_2,\ c).$$

So, the contribution of a given type of failure to the effect of combined injury may be reflected by the characteristics Π_1^l, Π_1^u and Π_2^l, Π_2^u along with the traditional ones $\Pi_1 = \mathrm{pr}(T_2 > T_1)$ and $\Pi_2 = \mathrm{pr}(T_1 > T_2)$. Using the same arguments as in the work of S. Rachev, A. Yakovlev [3] it is possible to generalize the above confidence bounds for the case of right-hand random censorship.

Numerical example. Consider the example presented in our previous paper (Rachev and Yakovlev). The same slight tendency to the prevalence of the second type of failure is revealed by use of interval estimators for the characteristics introduced in this communication. Actually the following confidence bounds with probability greater than 0.97 have been obtained:

$$-0.84 \leqq \Pi_2^l, \qquad \Pi_2^u \leqq 9.21$$

$$-1.09 \leqq \Pi_1^l, \qquad \Pi_1^u \leqq 8.16.$$

The example reaffirms the feeling that more restrictive assumptions concerning the structure of the competing risks model are necessary to strengthen the statistical inference from the survival of animals after combined injuries.

REFERENCES

1. M. Frecher. Sur les tableaux de correlation dont les margees sont donnes. *Ann. Univ. Lyon sect A, Ser. 3,* **14**, 1951, 53-77.
2. W. Hoeffding. Masstabinvariante Korrelationstheorie. *Schr. Math. Inst. Univ. Berlin,* 5, 1940, 181-233.
3. S. T. Rachev, A. Yu. Yakovlev. Bounds for the probabilistic characteristics of latent failure times within competing risks framework. *Serdica,* **14**, 1988.
4. D. Stoyan. Qualitative Eigenschaften und Abschätzungen stochastischer Modelle. Berlin, 1977.

SUNY at Stony Brook, N. Y. USA

Received 11. 2. 1987

Central Research Institute of Roentgenology and Radiology, Leningrad, USSR