Some Waiting Time Problems

Von L. RÅDE, Gothenburg[1])

1. Introduction

The classical waiting time problem is the following. Independent repetitions of an experiment with two outcomes A and B are performed until A occurs for the rth time. If this happens at trial number T, we call T the waiting time and if the probability for A is $P(A) = p = 1 - q$, we have

$$P(T = x) = \binom{x-1}{r-1} p^r q^{x-r}, \qquad x = r, \quad r+1, \ldots,$$

that is the waiting time T has a Pascal distribution. We call this problem the simple binomial waiting time problem.

Another wellknown waiting time problem is the following, which we call the simple hypergeometric waiting time problem. A set containing n elements, of which s, $1 \leqslant s \leqslant n$, are of a special kind, is sampled without replacement until an element of the special kind is found for the rth time, $1 \leqslant r \leqslant s$. If this happens for element number T taken from the set, T is the waiting time in this case. This problem and some other waiting problems are treated for instance in WILKS [1]. If in the last example the set is sampled with replacement, T has a Pascal distribution with $p = s/n$.

We will here first treat the simple hypergeometric waiting time problem in a partly new way deriving a generating function for the probability distribution of T. We will then introduce and solve a new type of waiting time problems originating from the sampling of a set of n elements but with the complication that the elements of the special kind only can be identified by comparing them with the elements of another set. We will call these problems double waiting time problems. They can serve as mathematical models for some psychological tests.

---

[1]) LENNART RÅDE, Chalmers Technical University, Gothenburg, Sweden.

2. The Simple Hypergeometric Waiting Time Problem

In this case the waiting time T has the sample space $\Omega_T = \{r, r+1, \ldots, n+r-s\}$. The event $\{T = k\}$ means that among the first $(k-1)$ elements there are $(r-1)$ elements taken from the set of the special kind and that one of the $(s-r+1)$ remaining elements of the special kind is then taken from the remaining $(n-k+1)$ elements.

We thus have

$$P(T = k) = p_k = \frac{\binom{s}{r-1}\binom{n-s}{k-r}}{\binom{n}{k-1}} \frac{s-r+1}{n-k+1}, \qquad k = r, r+1, \ldots, n+r-s$$

Writing the binomial coefficients in term of factorials we easily find that

$$P(T = k) = p_k = \frac{\binom{n-k}{s-r}\binom{k-1}{r-1}}{\binom{n}{s}}, \qquad k = r, r+1, \ldots, n+r-s \tag{1}$$

We will here use the probability generating function $\varphi(t) = \sum_k p_k t^{-k}$. The expectation μ and variance σ^2 is then given by

$$\mu = -\varphi'(1), \qquad \sigma^2 = \varphi''(1) + \varphi'(1) - \varphi'(1)^2 \tag{2}$$

We denote the probability generating function

$$\varphi_{s,n}(t) = \sum_{k=r}^{n+r-s} p_k t^{-k},$$

where p_k is given by (1). We can extend the summation to $k = n$ as $p_k = 0$ for $k = n+r-s+1, \ldots, n$. We thus have

$$(s-r)!\,(r-1)!\binom{n}{s}\varphi_{s,n}(t) = \tag{3}$$

$$\sum_{k=r}^{n} (n-k)(n-k-1) \ldots (n-k-s+r+1)(k-1) \ldots (k-r+1)\, t^{-k}$$

We now introduce the function

$$\Psi(t, u) = \sum_{k=r}^{n} t^{n-k} u^{k-1} = \frac{u^n - t^{n-r+1} u^{r-1}}{u-t}$$

By differentiation of $\Psi(t, u)$ with respect to t and u we find that the right hand member of (3) can be written

$$t^{s-n-r} \left[D_t^{(s-r)} D_u^{(r-1)} \{\Psi(t, u)\}\right]_{u=1}$$

where $D_t = \partial/\partial t$ and $D_u = \partial/\partial u$ are partial differentiation operators. We thus have

$$\varphi_{s,n}(t) = \frac{t^{s-n-r}}{(s-r)!\,(r-1)!\binom{n}{s}} [D_t^{(s-r)} D_u^{(r-1)}\{\Psi(t, u)\}]_{u=1}$$

If we now differentiate $\varphi_{s,n}$ with respect to t we find after some rearrangements that

$$\varphi'_{s,n}(t) = \frac{(s-r+1)\,(n-s)}{s+1} \varphi_{s+1,n}(t) + (s-n-r)\,\varphi_{s,n}(t)$$

If we here put $t = 1$ and observe that $\varphi_{s,n}(1) = 1$, we get from (2)

$$E[T] = \frac{r(n+1)}{s+1} \tag{4}$$

By differentiation once more we get

$$t\,\varphi''_{s,n}(t) = \frac{(s-r+1)\,(n-s)}{s+1} \varphi'_{s+1,n}(t) + (s-n-r-1)\,\varphi'_{s,n}(t)$$

From this we get according to (1) after some calculations

$$\text{Var}\,[T] = \frac{r(n+1)\,(n-s)\,(s-r+1)}{(s+1)^2\,(s+2)} \tag{5}$$

This variance is given in WILKS [1] under a more complicated form. When deriving the mean and variance we have supposed that $s < n$. The results arrived at however obviously hold for $s = n$.

We can now summarize in the following theorem.

Theorem 1. In the simple hypergeometric waiting time problem, the waiting time T has the probability generating function $\varphi_{s,n}$, where

$$\varphi_{s,n}(t) = \sum_{k=r}^{n+r-s} P(T=k)\,t^{-k} =$$

$$= \frac{t^{s-n-r}}{(s-r)!\,(r-1)!\binom{n}{s}} \left[\frac{\partial^{(s-r)}}{\partial t^{s-r}} \frac{\partial^{(r-1)}}{\partial u^{r-1}} \left\{\frac{u^n - t^{n-r+1}\,u^{r-1}}{u-t}\right\}\right]_{u=1}$$

the expectation

$$E[T] = \frac{r(n+1)}{s+1}$$

and the variance

$$\mathrm{Var}\,[T] = \frac{r(n+1)\,(n-s)\,(s-r+1)}{(s+1)^2\,(s+2)}$$

3. Double Waiting Time Problems

By a double waiting time problem we mean a problem with the following structure. A set of n_1 elements contains s_1, $1 \leqslant s_1 \leqslant n_1$, elements of a special kind, which can be identified only by comparing them to the elements in another set. In this other set there are n_2 elements but only s_2, $1 \leqslant s_2 \leqslant n_2$, of these elements are such that they will identify the s_1 elements of the special kind in the first set.

The elements in the first set can be thought of as keys and the elements in the other set as locks. The elements of the special kind in the first set are then s_1 keys only fitting s_2 locks in the set of locks.

We now suppose that a man tries to find a key fitting one of the locks in such a way that he takes a key at random and tries it with successive locks selected at random and without replacement from the set of locks. If a chosen key does not fit any one of the locks, he takes a new key and repeats the same procedure. The waiting time T is the number of that trial, when for the first time he has a key fitting a lock. We suppose that the number s_2 is not known to the man. Thus he always has to compare a chosen key, which is not of the special kind with all the locks. If s_2 were known to the man, he could always take a new key, if $(n_2 - s_2 + 1)$ trials with one key have been unsuccessful.

If the keys are taken at random with replacement we call the problem the double binomial waiting time problem and if the keys are taken at random without replacement we call the problem the double hypergeometric waiting time problem.

Some psychological testing procedures are of this structure, for instance testing procedures for testing men's mechanical ability. The model corresponds to the null hypothesis that a person has no ability at all and chooses keys and locks at random.

4. The Double Binomial Waiting Time Problem

The sample space Ω_T of the waiting time T is in this case Ω_T, where

$$\Omega_T = \{1, 2, \ldots, n_2 - s_2 + 1\} \cup \{n_2 + 1, n_2 + 2, \ldots, 2n_2 - s_2 + 1\} \cup \ldots =$$

$$= \bigcup_{k=1}^{\infty} \{x \in N \mid 1 + (k-1)\,n_2 \leqslant x \leqslant 1 + k\,n_2 - s_2\}.$$

Here N is the set of natural numbers.

We now regard the event $\{T = p + (k-1)\, n_2\}$, $1 \leq p \leq n_2 - s_2 + 1$, which means that key number k taken from the first set is the first chosen key of the special kind and that lock number p is the first chosen lock fitting this key. We thus have

$$P(T = (k-1)\, n_2 + p) = p\, q^{k-1} \frac{\binom{n_2 - p}{s_2 - 1}}{\binom{n_2}{s_2}}$$

The generating function φ_T for this probability distribution is now given by

$$\varphi_T(t) = \sum_{k=1}^{\infty} \sum_{p=1}^{n_2 - s_2 + 1} P(T = (k-1)\, n_2 + p)\, t^{-(k-1)n_2 - p} =$$

$$= \sum_{k=1}^{\infty} p\, q^{k-1}\, t^{-(k-1)n_2} \sum_{p=1}^{n_2 - s_2 + 1} \frac{\binom{n_2 - p}{s_2 - 1}}{\binom{n_2}{s_2}}\, t^{-p} =$$

$$= t^{n_2} \sum_{k=1}^{\infty} p\, q^{k-1}\, (t^{n_2})^{-k}\, \varphi_{s_2, n_2}(t) = \varphi_{s_2, n_2}(t)\, t^{n_2}\, \varphi_p(t^{n_2})$$

Here is $\varphi_p(t) = p/(t-q)$, $q = 1 - p$, the generating function for the waiting time in the binomial simple waiting time problem with $r = 1$. Further we denote here and in the following by $\varphi_{s,n}$ the generating function for the waiting time in the simple hypergeometric waiting time problem with parameters s, n and $r = 1$. $\varphi_{s,n}$ is given in theorem 1 above.

We can now get the expectation and variance of the waiting time by differentiation of the generating function. We observe that the generating function for a sum of independent stochastic variables is the product of the generating functions for the variables and that if the stochastic variable ξ has the generating function φ, then $t^n \varphi(t^n)$ is the generating function for $n(\xi - 1)$. We obviously have that the waiting time in this example can be written $T = T_2 + \eta_2(T_1 - 1)$, where T_1 and T_2 are the waiting times in the simple binomial and hypergeometric waiting time problem respectively.

That the waiting time in this case has this structure can be seen directly in the following way. If T_1 is the waiting time for a key of the special kind, $n_2(T_1 - 1)$ trials have to be made with the wrong keys and then T_2 comparisons have to be made with the first key of the special kind.

We now have

$$E[T] = E[T_2] + n_2(E[T_1] - 1) = \frac{n_2 + 1}{s_2 + 1} + \frac{q}{p} n_2$$

$$\mathrm{Var}[T] = \mathrm{Var}[T_2] + n_2^2 \,\mathrm{Var}[T_1] = \frac{s_2(n_2 + 1)(n_2 - s_2)}{(s_2 + 1)^2 (s_2 + 2)} + \frac{q}{p^2} n_2^2$$

We have thus proved the following theorem.

Theorem 2. In the double binomial waiting time problem the waiting time T has the generating function $\varphi_{s_2, n_2}(t)\, t^{n_2} \varphi_p(t^{n_2})$, the expectation

$$E[T] = \frac{n_2 + 1}{s_2 + 1} + \frac{q}{p} n_2$$

and the variance

$$\mathrm{Var}[T] = \frac{s_2(n_2 + 1)(n_2 - s_2)}{(s_2 + 1)^2 (s_2 + 2)} + \frac{q}{p} n_2^2$$

5. The Double Hypergeometric Waiting Time Problem

In this case the sample space of the waiting time T is

$$\Omega_T = \bigcup_{k=1}^{n_1 - s_1 + 1} \{x \in N \mid 1 + (k - 1)\, n_2 \leqslant x \leqslant 1 + k\, n_2 - s_2\}.$$

In the same way as in the binomial case we derive the relation $T = T_2 + n_2(T_1 - 1)$, where T_1 and T_2 are independent waiting times for the simple hypergeometric waiting time problems with parameters s_1, n_1 and s_2, n_2 respectively. We thus immediately get the following theorem.

Theorem 3. In the double hypergeometric waiting time problem the waiting time T has the generating function $\varphi_{s_2, n_2}\, t^{n_2} \varphi_{s_1, n_1}(t^{n_2})$, the expectation

$$E[T] = \frac{n_2 + 1}{s_2 + 1} + n_2 \frac{n_1 - s_1}{s_1 + 1}$$

and the variance

$$\mathrm{Var}[T] = \frac{s_2(n_2 + 1)(n_2 - s_2)}{(s_2 + 1)^2 (s_2 + 2)} + n_2^2 \frac{s_1(n_1 + 1)(n_1 - s_1)}{(s_1 + 1)^2 (s_1 + 2)}$$

References

[1] Wilks, S. S.: Mathematical Statistics, New York, 1962. 141–144.