The Simple Model of the Recurrence Time in the Reproductive Pattern of a Married Female

CHI-HOON CHOI

0. Introduction

This paper is to treat entirely through the Bayesian Statistics. The recurrence time has been studied by many scholars but this shall be the first trial to approach by the subjective probabilistic idea.

Biological considerations make it natural to view as a random variable the length of time spent in a state after the beginning of marriage or after return from another state.

The model which is used in this paper is very simple one. We may introduce more complicated states such as not only a live-birth but also still-birth and induced or spontaneous abortion. However, the Bayesian approach can also be handled with stochastic process for this problem.

This paper does not include how to decide the parameters which appear in prior distribution and how to collect data to fit formulas.

1. The model

One sets up the model of the reproductive pattern of a married female under the assumption that a woman is in one and only one of the following states;

s_0 = nonpregnant and fecundable,

s_1 = pregnant state to be terminated in livebirth,

s_2 = postpartum sterile period associated with only livebirth.

We shall assume that a female in the state s_0 is subject at intervals of approximately one month to a probability $\tilde{p}_1$ of entering state s_1, that the female in the state s_1 has the probability 1 of entering state s_2, and that the female in the state s_2 is subject to a probability $\tilde{p}_2$ of entering state s_0.

Then the total number of months t_1 spent in state s_0 during any single visit to that state has a density

$$g_1(t_1|\tilde{p}_1) = \tilde{p}_1(1-\tilde{p}_1)^{t_1-1}, \quad t_1 = 1, 2, \cdots\cdots; 0 \le \tilde{p}_1 \le 1.$$

Similarly we have for state s_2

$$g_2(t_2|\tilde{p}_2) = \tilde{p}_2(1-\tilde{p}_2)^{t_2-1}, \quad t_2 = 1, 2, \cdots\cdots; 0 \le \tilde{p}_2 \le 1,$$

Received by the editors July 20, 1968.

where t_2 is the total number of months spent in state s_2 during a single visit to that state.

But in state s_1, the time t spent in that state is to be assumed that t' is distributed by a Normal density with a known and an unknown σ^2.

Now we can see very simple reproductive process in which a woman, beginning in state s_0, passes after a period of time into state s_1, thence, with a probability s_2 and back to state s_0, to begin the reproductive cycle again. This is a Renewal Process.

2. The likelihood and sufficient statistics

(i) *For the states s_0 and s_2*

The likelihood of t_{1i} is

$$g_1(t_{1i}|p_1)=p_1(1-p_1)^{t_{1i}-1}$$

thus the likelihood of observations $t_{11}, \cdots\cdots, t_{1n}$, which are independently and identically distributed(*iid*) by $t_{1i}\sim p_1(1-p_1)^{t_{1i}-1}$, given p_1 is given by

$$g_1(t_{11}, \cdots\cdots, t_{1n}|p_1)\propto \prod_{i=1}^{n} p_1(1-p_1)^{t_{1i}-1}=p_1^{n}(1-p_1)^{\sum_1^n t_{1i}-n}.$$

Therefore $(n, \sum_1^n t_{1i})$ is sufficient for p_1.

(ii) *For state s_1*

The likelihood of t' is

$$g(t_1', \cdots\cdots, t_m'|\mu, \sigma)\propto \frac{1}{\sigma^m}\exp\left\{-\frac{\sum(t_i'-\mu)^2}{2\sigma^2}\right\}\propto\frac{1}{\sigma^m}\exp\left\{-\frac{ms^2}{2\sigma^2}\right\}$$

where $s^2\equiv\frac{\sum(t_i'-\mu)^2}{m}$ and m is a number of observations which are *iid* a Normal with a mean μ and a variance σ^2.

Thus the sufficient statistics are m and s^2 for σ^2 when μ is known.

3. The conjugate prior and the posterior distribution

(i) *For the states s_0 and s_2*

(a) *The conjugate prior*

The conjugate prior for p_1 is appropriate if it is Beta density, that is,

$$g_1'(p_1|v_0, v_1)=k\cdot p_1^{v_0-1}(1-p_1)^{v_1-v_0-1},\ \ 0\le p_1\le 1;\ v_1>v_0>0,$$

where $k=\frac{(v_1-1)!}{(v_0-1)!(v_1-v_0-1)!}$ and v_0, v_1 are determined by prior knowledges such that

$$\mu_1=v_0/v_1,\quad \mu_2=\frac{v_0(v_1-v_0)}{v_1^2(v_1+1)}$$

with $v_1 > v_0 > 0$.

(b) *The posterior density of* $\tilde{p}_1$

The posterior density of $\tilde{p}_1$ is proportional to the product of the likelihood and the prior, i.e.,

$$g''(p_1 | t_{11}, t_{12}, \cdots\cdots, t_{1n}) \propto p_1{}^n (1-p_1)^{\sum_1^n t_{1i}-n} \cdot p_1{}^{\nu_0-1}(1-p_1)^{\nu_1-\nu_0-1}.$$

The standardized constant will be obtained by the integration

$$\int_0^1 p_1{}^{n+\nu_0-1}(1-p_1)^{n(\bar{t}_1-1)+\nu_1+\nu_0-1} dp_1 = \frac{(n+v_0-1)!(n\bar{t}_1-n+v_1-v_0-1)!}{(n\bar{t}_1+v_1-1)!}$$

where
$$\bar{t}_1 = \frac{1}{n}\sum_i^n t_{1i}.$$

Hence the posterior density of $\tilde{p}_1$ given the data and v_0, v_1 is

$$g''(p_1 | \text{data}, v_0, v_1) = \frac{(n\bar{t}_1+v_1-1)!}{(n+v_0-1)!(n\bar{t}_1-n+v_1-v_0-1)!} p_1{}^{n+\nu_0-1}(1-p_1)^{n(\bar{t}_1-1)+\nu_1-\nu_0-1},$$
$$0 \leq p_1 \leq 1;\ v_1 > v_0 > 0.$$

(c) *Unconditional density of* t_1

The unconditional density of t_1 can be obtained by

$$\begin{aligned} g_1(t_1 | v_0, v_1) &= \int_0^1 g_1(t_1 | p_1) g'(p_1 | v_0, v_1) dp_1 \\ &= \int_0^1 p_1(1-p_1)^{t_1-1} \frac{(v_1-1)!}{(v_0-1)!(v_1-v_0-1)!} p_1{}^{\nu_0-1}(1-p_1)^{\nu_1-\nu_0-1} dp_1 \\ &= \frac{(v_1-1)!v_0!(t_1+v_1-v_0-2)!}{(v_0-1)!(v_1-v_0-1)!(t_1+v_1-1)!}. \end{aligned}$$

Similarly we can obtain that

$$g_1(\bar{t}_1 | v_0, v_1;\ n) = \frac{(v_1-1)!(v_0+n-1)!(n\bar{t}_1-nv_1-v_0-1)!}{(v_0-1)!(v_1-v_0-1)!(n\bar{t}_1-v_1-1)!}.$$

(ii) *For the state* s_1

(a) *Conjugate prior*

When the mean μ of an independent normal process is known but the precision σ^2 is treated as a random variable, the natural conjugate of

$$\frac{1}{\sigma^m} \exp\left\{-\frac{ms^2}{2\sigma^2}\right\}$$

is

$$g'(\sigma^2 | v_0, v_1) = K\sigma^{-(v_0+1)} \exp\left\{-\frac{v_0 v_1}{2\sigma^2}\right\}$$

where $K=\dfrac{(v_0v_1)^{v_0/2}}{2^{\frac{v_0-2}{2}}\Gamma\left(\frac{v_0}{2}\right)}$ this is known as the Gamma-2 distribution.

(b) *Unconditional distribution of* t'

$$g(t'|v_0,\ v_1,\ \mu)=\int_0^\infty g(t'|\mu,\sigma^2)g'(\sigma^2|v_0,\ v_1)d\sigma^2$$

$$=\int_0^\infty \frac{K}{\sqrt{2\pi}\,\sigma}e^{-\frac{(t'-\mu)^2}{2\sigma^2}}\cdot\frac{1}{\sigma^{v_0+1}}e^{-\frac{v_0v_1}{2\sigma^2}}d\sigma^2\propto[(t'-\mu)^2+v_0v_1]^{-\frac{v_0+1}{2}}$$

$$\propto\left[1+\frac{(t'-\mu)^2}{v_0v_1}\right]^{-\frac{v_0+1}{2}}.$$

Thus $\dfrac{t'-\mu}{\sqrt{v_1}}$ has t-distribution with v_0 degrees of freedom, i.e., if we put $T=\dfrac{t'-\mu}{\sqrt{v_1}}$ then

$$g(T|v_0,\ v_1,\ \mu)=\frac{\Gamma\left(\frac{v_0+1}{2}\right)}{\sqrt{\pi}\,v_0\Gamma\left(\frac{v_0}{2}\right)}\ \frac{1}{\left[1+\frac{T^2}{v_0}\right]^{\frac{v_0+1}{2}}}.$$

4. The expectation of t and t'

(i) t_1 has the density

$$g_1(t_1|v_0,\ v_1)=\frac{(v_1-1)!\,v_0!(t_1+v_1-v_0-2)!}{(v_0-1)!(v_1-v_0-1)!(t_1+v_1-1)!}$$

$$E(t_1)=\sum_{t_1=1}^{\infty}\frac{(v_1-1)!\,v_0!(t_1+v_1-v_0-2)!}{(v_0-1)!(v_1-v_0-1)!(t_1+v_1-1)!}\,t_1$$

$$=\sum_{t_1=1}^{\infty}\int_0^\infty t_1\frac{(v_1-1)!}{(v_0-1)!(v_1-v_0-1)!}\,p_1^{v_0}(1-p_1)^{t_1+v_1-v_0-2}dp_1$$

$$=\int_0^\infty\frac{(v_1-)!}{(v_0-1)!(v_1-v_0-1)!}\,p_1^{v_0-1}(1-p_1)^{v_1-v_0-2}\sum_{t_1=1}^{\infty}t_1p_1(1-p_1)^{t_1}dp_1$$

$$=\frac{(v_1-1)!}{(v_0-1)!(v_1-v_0-1)!}\int_0^\infty p_1^{v_0-1}(1-p_1)^{v_1-v_0-2}\frac{1-p_1}{v_1}\,dp_1$$

$$=\frac{(v_1-1)!}{(v_0-1)!(v_1-v_0-1)!}\cdot\frac{(v_0-2)!(v_1-v_0-1)!}{(v_1-2)!}$$

$$=\frac{v_1-1}{v^0-1}$$

(ii) $T=\dfrac{t'-\mu}{\sqrt{v_1}}$ has the density of Student. Hence $E(T)=0$, that is $E(t')=\mu$.

5. Recurrence time

We can see that t_1, t', t_2 are independent. Therefore the likelihood of t_1, t', t_2

given p_1, μ, σ^2, p_2 is

$$f_l(t_1,\ t',\ t_2 | p_1,\ \mu,\ \sigma^2,\ p_2) = p_1(1-p_1)^{t_1-1} \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(t'-\mu)^2}{2\sigma^2}} p_2(1-p_2)^{t_2-1}$$

and the prior of p_1, σ^2, and p_2 is

$$f_P(p_1,\ \sigma,\ p_2 | v_0,\ v_1\,;\ v_0,\ v_1\,;\ v_2,\ v_3)$$
$$\propto p_1^{v_0-1}(1-p_1)^{v_1-v_0-1}\,\sigma^{-(v_0+1)}\,e^{-\frac{v_0 v_1}{2\sigma^2}}\,p_2^{v_2-1}(1-p_2)^{v_3-v_2-1}.$$

Hence unconditional joint density of t_1, t', t_2 should be

$$f(t_1,\ t',\ t_2 | v_0,\ v_1,\ v_0,\ v_1,\ v_2,\ v_3)$$
$$\propto \frac{(t_1+v_1-v_0-2)!\,(t_2+v_2-v_3-2)!}{(t_1+v_1-1)!\,(t_2+v_3-1)!}\left[1+\frac{(t'-\mu)^2}{v_0 v_1}\right]^{-\frac{v_0+1}{2}}$$

Let $t_1+t'+t_2=t$ and let R_t denote the set of all t_1, t', t_2 which satisfy $t_1+t'+t_2 \leqq t$, then the probability that the recurrence time is less than or equal to a given t_0 is given by

$$P(t_1+t'+t_2 \leqq t_0) = \sum_{t_1}\int_{t'}\sum_{t_2} f(t_1,\ t',\ t_2 | v_0,\ v_1,\ v_0,\ v_1,\ v_2,\ v_3),$$
$$t_1,\ t',\ t_2 \in R_{t_0}.$$

6. The assumptions on p_1, σ^2 and p_2

p_1 is called by fecundability in common. The fecundability has been estimated by quite a few scholars under various models. Typical models are as follows:

(i) *The model based on the length of the fertile period and the frequency of coitus per menstrual cycle.*

(ii) *The model using of the total number of children born to couples over* 45 *in a population reputed to completely avoid contraception.*

(iii) *The model fitting a theoretical distribution to experience respecting first conceptive delays.*

We have been considering third model, thus the assumption on the fecundability p_1 of each couple has to remain p_1 constant from month to month until conception.

If p_1 or p_2 are not identical for each couple the likelihood of t_{1i} will be

$$g_1(t_{1i} | p_{1i}) = p_{1i}(1-p_{1i})^{t_{1i}-1}$$

thus for t_{11}, t_{12}, ……, t_{1n}

$$g_1(t_{11},\ t_{12},\ \cdots\cdots,\ t_{1n} | p_{11},\ \cdots\cdots,\ p_{1n}) \propto \prod_{i=1}^{n} p_{1i}(1-p_{1i})^{t_{1i}-1}.$$

This can be written as

$$g_1(\underline{t}_1 | \underline{p}_1) \propto \exp\left\{\sum_1^n \ln p_{1i} + \sum_1^n (t_{1i}-1)\ln(1-p_{1i})\right\}$$
$$\propto G(\underline{p}_1) \exp\left\{\sum_1^n f(t_{1i})\phi(p_{1i})\right\}$$

Note that in fact there may not be all different $p_{11}, \cdots\cdots, p_{1n}$, but still we may say that, therefore, $(t_{11}, \cdots\cdots, t_{1n}, n)$ is jointly sufficient for $p_{11}, \cdots\cdots, p_{1n}$.

The prior of $p_{11}, \cdots\cdots, p_{1m}$ should be

$$g'(p_{11}, \cdots\cdots, p_{1m}) \propto p_{11}^{\nu_{01}-1}(1-p_{11})^{\nu_{11}-\nu_{01}-1} \cdots\cdots p_{1m}^{\nu_{0m}-1}(1-p_{1m})^{\nu_{1m}-\nu_{0m}-1}$$

where m is the number of groups which we will assume it known and categorized by some characteristic such as age at marriage or parity if p_1 is replaced in p_2.

Then the posterior would be

$$g''(p_{11}, p_{22}, \cdots\cdots, p_{1m} | \underline{t}_{11}, \cdots\cdots, \underline{t}_{1m})$$
$$\propto p_{11}^{n_1+\nu_{01}-1}(1-p_{11})^{n_1(t_{11}-1)+\nu_{11}-\nu_{01}-1} \cdots\cdots p_{1m}^{n_m+\nu_{01}-1}(1-p_{1m})^{n_m(t_{1m}-1)+\nu_{1m}-\nu_{0m}-1}$$

where $n_1+n_2+\cdots\cdots+n_m = n$, all n_i's are known.

This result can be easily derived by the independent of t_k's. Hence the unconditional distribution of $t_{11}, \cdots\cdots, t_{1m}$ can be expressed by

$$g_1(t_{11}, \cdots\cdots, t_{1m} | \nu_{01}, \nu_{02}, \cdots\cdots, \nu_{0m}, \nu_{11}, \nu_{12}, \cdots\cdots, \nu_{1m})$$
$$= \prod_{j=1}^m \frac{(\nu_{1j}-1)!(\nu_{0j}-n_j-1)!(n_j\bar{t}_{1j}-n_j\nu_{1j}-\nu_{0j}-1)!}{(\nu_{0j}-1)!(\nu_{1j}-\nu_{0j}-1)(n_j\bar{t}_{1j}-\nu_{1j}-1)!} .$$

Bibliography

1. Hartman, C.G., *Science and the Safe Period,* Williams and Wilkins, Baltimore.
2. Lindley, D.V., *Introduction to Probability and Statistics from a Bayesian* Viewpoint, Part 1 and 2, Cambridge Univ. Press, 1965.
3. Perrin, E.B. and Sheps, M.C., *Human Reproduction: A Stochastic Process,* Biometrics **20,** No. 1, pp. 28-45.
4. Raiffa, H. and Schlaifer, R., *Applied Statistical Decision Theory,* Division of Research, Graduate School of Business Administration, Harvard University, 1961.

Seoul National University
University of Michigan