International Journal of Statistics and Probability; Vol. 5, No. 2; 2016
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

# On Predicting Survival in Prostate Cancer: Using an Extended Maximum Spacing Method at the Change Point of the Semiparametric Ratio Estimator (SPRE)

Deborah Weissman-Miller[1]

[1] School of Occupational Therapy, College of Health Sciences, Brenau University, Gainesville, GA USA

Correspondence: Deborah Weissman-Miller, School of Occupational Therapy, College of Health Sciences, Brenau University, Gainesville, GA USA. Tel: 850-570-7696. E-mail: dweissman-miller@brenau.edu

Received: January 5, 2016 Accepted: January 22, 2016 Online Published: February 10, 2016

doi:10.5539/ijsp.v5n2p19 URL: http://dx.doi.org/10.5539/ijsp.v5n2p19

**Abstract**

Prostate cancer is a condition of public health significance in the United States. A new method for predicting survival is derived for the domain around the change point from a semiparametric ratio estimator (SPRE) to predict survival in response to treatment for prostate cancer. Using an extended maximum spacing estimator, the geometric mean of sample spacings from a uniform distribution $U(u,v)$ is derived with known endpoints given at 0 and at the value of the change point from an ordinary least squares (OLS) regression for SPRE. To determine the maximum interval on the 'x' axis between point estimates, the maximum spacing estimation method is derived from a continuous univariate distribution where spacing will be defined as gaps between ordered values of the distribution function. The maximum is defined as a single value in the neighborhood of the change point and spacing defined as a function of time. This maximum spacing defines the gaps between point estimates at each time-dependent predicted outcome from the change point and results in a semiparametric ratio estimator that is reliable and repeatable. Performance is discussed through a simulation of change point values for a real application in clinical medicine and, using SPRE, in personalized medicine for a single prostate cancer patient.

**Keywords: s**urvival predictions, prostate cancer, maximum spacing estimation, SPRE

## 1. Introduction

### *1.1 Introduction to Prostate Cancer Survival Prediction*

The motivation for this paper is to accurately predict survival in response to treatment for prostate cancer using a novel extended maximum spacing derivation to determine maximum time intervals between point estimates for a new semiparametric ratio estimator (SPRE) from Weissman-Miller (2013). Maximum spacing is particularly important in predicting PSA levels or Gleason scores over time either in "watchful waiting" or in survival after treatment in prostate cancer. Since prostate cancer has a long natural history, accurate survival predictions may be important in determining treatment effectiveness. Prostate cancer is the most prevalent cancer of adult men in the United States (American Cancer Society, 2009) and Western Europe (Eton & Lepore, 2002), and the second most common cause of cancer death after lung cancer according to Pastides (2001). The National Cancer Institute estimates that there were 192,280 new cases of prostate cancer identified in the United States during 2009, and that 27,360 men will die of this cancer (National Cancer Institute, 2009). There are a number of treatment modalities currently available for prostate cancer according to Pastides (2001), with new variations being tested in clinical trials all the time as noted by Eggener & Coleman (2008) and Tempany, Straus, Nabuhiko & Haker, (2008). Options range from 'watchful waiting' at the extreme low end of the intensity spectrum, to radical prostatectomy with orchiectomy or androgen deprivation and/or other chemotherapy, used for advanced disease, at the upper end of the intensity spectrum. Each treatment in this range has possible benefits, adverse effects and medical costs associated with it. The use of the SPRE model to predict patient survival after treatment can enhance the use of differing treatments that may optimize predicted survival. Since SPRE functions well as a single-subject design model as shown in Weissman-Miller, Shotwell and Miller (2012) and in Weissman-Miller (2013)], SPRE can function as a 'Personalized Medicine' model for survival prediction in prostate cancer.

### *1.2 Explore an Original Prostate Cancer Study*

An examination of the data reported by Joseph, Al-Qaisieh, Ash, Bottomley & Carey (2004) indicates that there is an

apparent initial linearity from the onset of measuring data for freedom from biochemical failure in patients with a biopsy Gleason score equal to 7, and less than or greater than 7 as well. The dataset predicts the relapse-free survival of 667 patients with localized prostate cancer treated by brachytherapy with implantation of $I^{125}$ seeds as monotherapy. This is one of the largest series of such patients reported thus far. Actuarial survival curves were initially calculated by the Kaplan-Meier method. Cox proportional-hazards multivariate analysis was used to ssess the influence of covariates on the reported results from Joseph, et al (2004).

*1.3 Introduction to a Novel Prostate Cancer Survival Prediction*

Long-term patient outcomes are predicted from short-term linear regression of initial patient's data from a change point for a novel semiparametric ratio estimator (SPRE). The change point is the session number or time interval where the patient adapts to the treatment according to Weissman-Miller, et al (2012). Then a novel extended maximum spacing estimate in the domain of the change point is derived to predict outcome spacing from the predictive distribution values relative to clinical test data. These estimates are on brachytherapy for prostate cancer from a survival analysis where the primary test data results are analysed using a Kaplan-Meier non-parametric method. The intent of this paper is to determine the maximum spacing of the predicted outcomes from the change point at 23 months to the $89^{th}$ month after treatment. The results relate very well to the original survival analysis data with an error of +/- 1.15% at 89 months.

## 2. Method

*2.1 The Statistical Method of the Change Point and Response Function in SPRE for Maximum Spacing*

The extended maximum spacing is derived around the domain of the change point in the region from time 0 to the time at the change point of 23 months for Gleason score = 7 (Joseph et al, 2004). In the SPRE model, the change point is derived from a backwards stepwise ordinary least squares regression, which provides minimum bias. The change point is determined from the highest or lowest F statistic given as $F = t^2 = MS\,\mathrm{Re}\,g / \hat{\sigma}^2$ that is associated with the relevant distributional P-value (Weissman-Miller, 2013) at the value of the time interval in prostate cancer or session number (in many therapies). From the change point, nonlinear point estimates are given by SPRE model using a new response function $G_k, \tau(t)$ given by Weissman-Miller (2013) from the cumulative distribution function (CDF) of the Weibull distribution. The point estimates are given from a ratio of this function times the prior estimated outcome, $\hat{\theta}_t = \frac{G_k, \tau(t_{i+1})}{G_k, \tau(t_i)} \cdot \theta_{t_i}$ (Weissman-Miller, 2013). The step function is given here as $t_{i+1}$ for maximum spacing. It should be noted that the SPRE estimator is consistent at the beginning and end points of analysis, the upper and lower bounds, because it is unbiased at the change point and when the ratio $R = 1.00$. Pilot studies for 13 – 14 total therapy sessions where there is a low value of the change point, have shown an excellent relation of predicted to test data using $t_{i+1}$ in SPRE in occupational therapy for such areas as fall prevention in adults (Weissman-Miller & Graham, 2015). A preliminary study for point estimations in prostate cancer survival from a higher value change point using varying maximum spacing was presented in Weissman-Miller (2011). The importance of deriving accurate survival point estimations for prostate cancer cannot be overstated as prostate cancer has a long natural history and a timely prediction of survival may enable a tailored approach to prevention and care.

*2.2 New Predictions for Maximum Spacing Estimation for the Model Data*

The question now is – for predicting survival outcomes in prostate cancer, what is the statistically derived maximum spacing to use for point estimations from SPRE for a longer-term change point?

The maximum spacing estimation (MSE) in this paper is derived from the ordinary least squares (OLS) line from 0 to the value of the change point for the data at 23, in months, after which nonlinear point estimates are given by SPRE model. Since the change point is at some distance along the 'x' axis from 0, the ratio of Weibull distributions as given by Weissman-Miller, et al (2012) and Weissman-Miller (2013) is shifted so that the maximum likelihood methods will fail. Therefore, an extended maximum spacing method is utilized in this paper to derive the equations for $U(u,v)$ from which a value of $\tilde{v}$ is specified as a reduced value, and is derived with a complementary value to the

originally derived equations. Then the maximum point estimate and spacing gap can be derived for the 'x' axis. The original maximum spacing method used in this paper is based on the work by Cheng and Amin (1983) and Ranneby (1984), with an extension to the method by Cheng and Stephens (1989). In Ranneby's method, an approximation to the Kullback-Leibler information is obtained by using spacing where each component is bounded from above. In general, the concepts underlying the MSE method are based on the probability integral transform at the 'true parameter', where the 'spacing' between each observation should be uniformly distributed. In this case, the implication is that the difference between the values of the cumulative distribution function at consecutive observations should be equal. This methodology has profound consequences for predicting outcomes of PSA or, in this paper, Gleason scores in prostate cancer. If the equal spacing of point estimates can be predicted over time using this method that maximizes the geometric mean of the spacing, then the 'best fit' for these spacings is obtained. Using the following extended maximum spacing methods will ensure that the most reliable survival prediction can be made from the change point. A maximum spacing is derived on the 'x' axis for predictions from the change point, as shown on Figure (1) as the distance between $\hat{x}_{(n)}$ and $x_{(n)}$.

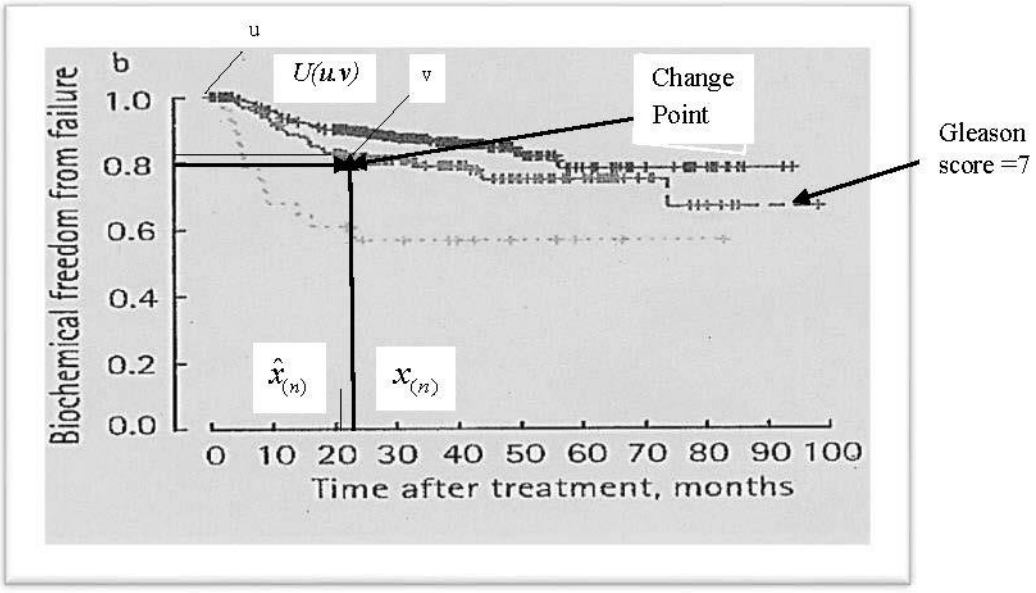


Figure 1. The change point (23 months) on the OLS, *U(u,v)* and the spacing $\hat{x}_{(n)} - x_{(n)}$.

(Permission granted: Joseph, et al (2004). Prostate-specific antigen relapse-free survival in patients with localized prostate cancer treated by brachytherapy. *BJU International; 94)*

*2.3 Definition of Spacings as* $D_i(\theta)$

By spacings, we refer to the gaps (distances) between successive points on a line. The estimation of parameters is applied in any continuous univariate distribution. Following Cheng and Amin (1983), this estimation of parameters is applied to a distribution with density $f(x,\Theta)$ and CDF, $F(x,\Theta)$. For a univariate distribution, let $\{x_{(1)},...,x_{(n)}\}$ be an ordered random sample of size 'n' drawn from the distribution from smallest to largest. The initial spacings are the gaps between values of point estimations of the distribution function at adjacent ordered points. Then, following Cheng and Amin (1983), Wong and Li (2006) and Ghosh & Jammalamadaka (2001):

$$D_i(\theta) = F_\theta(x_{(i)}) - F_\theta(x_{(i-1)}),\ \ i = 1,...,n+1 \tag{1}$$

The original spacing function for $U(u,v)$ may be given by:

$$S_n(\theta) = \ln \sqrt[n+1]{D_1 D_2 ... D_{n+1}} \tag{2}$$

In this paper, the following derivation of the univariate case is given from equation (2):

$$S_n(\theta) = \frac{1}{n+1}\sum_{i=1}^{n+1} \ln D_i(\theta) \tag{3}$$

It should be noted that the analysis in this paper is aimed at defining the spacing on the "x" axis from the maximum spacing equations used to define the initial parameters. In this case, the original derivation should maximize the geometric mean of the spacing, where the geometric mean is given as the $n^{th}$ root of their product. Then the maximum spacing estimator of $\theta_0$ is a value that maximizes the logarithm of the geometric mean of the sample spacing. For this analysis, the logarithm is given as the natural log, ln, which is often used in the spacing estimates for statistical entropy. In this sense, the message (in information theory) stands for an event, sample or character drawn from a distribution or data stream. In this analysis, any uncertainty characterized would relate to the disease process.

*2.4 Derivation of the MSE and the Maximum Spacing of a Uniform Distribution*

The maximum spacing estimator (MSE) of $\theta_0$ is given as the argmax of the spacing function, where the argmax is defined as a point:

$$\hat{\theta} = \arg\max S_n(\theta) \tag{4}$$

The maximum spacing is defined along this linear curve at the change point for the values of the Gleason score = 7 in the prostate cancer data that lie along the linear least squares regression line. The cumulative distribution function of the continuous uniform distribution is given as:

$$F(x) = \begin{cases} 0 & for\ x < u \\ \dfrac{x-u}{v-u} & for\ x \in [u,v] \\ 1 & for\ x \geq v \end{cases} \tag{5}$$

Then from equation (5), each discrete spacing is given by:

$$D_1 = \frac{x_{(1)} - u}{v-u}, \quad D_i = \frac{x_{(i)} - x_{(i-1)}}{v-u} \quad \text{for i =2,…,n} \tag{6}$$

And:

$$D_{n+1} = \frac{v - \tilde{x}_{(n,\tilde{v})}}{v-u} \tag{7}$$

2.4.1 Derivation for Parameters for the Extended Maximum Spacing Solution

In this extended analysis, the assumption is made that the values between spacing on the OLS line of values between 0 and 23 months are linear and assumed to be finite, and the additional assumption is made that the spacing between point estimations on the values of time for the Weibull ratio function are linear. The continuous uniform distribution $U(u,v)$ fits these assumptions, such that for each member of the family all intervals of the same length are equally probable. The support is defined by the two parameters $u$ and $v$, which are defined as the minimum and maximum values of the OLS line from 0 to the change point. Therefore, from values on the OLS line, an assumption is made that the maximum spacing estimates $U(u,v)$ are known. For survival analysis, the extended derivation includes an unknown parameter (where a value of the initial parameter $\tilde{v}$ is specified as a reduced value of the initial parameter) and given in this analysis as a function of $x_{(n)}$. Then $x_{(n)}$ is initially defined as $\tilde{x}_{(n,\tilde{v})}$, the reduced value of the $x_{(n)}$ variable expressed in terms of the $v$ scale. However, the values in survival analysis descend from 1 to 0. Then the reduced value of $\tilde{v}$ will be smaller than $v$. Thus, this value will be below the change point, as defined in Figure (1). However, the final transposed value of $\tilde{x}_{(n,\tilde{v})}$ is within the support of the $U(u,v)$ distribution on the 'x' axis. Finally, $\hat{x}_{(n)}$ as a point estimate of the 'x' axis (lower bound of maximum spacing) is derived from $\tilde{x}_{(n,\tilde{v})}$. This makes sense

when visualizing the relationship between $x_{(n)}$ =23 on the x axis, when n=23 months at the change point, corresponding to $v$ =0.84 on the y axis, where $\tilde{x}_{(n,\tilde{v})}$ is outlined below the value of *v* and the derived $\hat{x}_{(n)}$ from $\tilde{x}_{(n,\tilde{v})}$ is drawn near the value of $x_{(n)}$ at the change point. This is true when the values are related by the OLS line and the continuous uniform distribution modeled from it as $U(u,v)$ in Figure (1).

In this analysis, the region $D_{n+1}$ includes a new subscript for $x_{(n)}$ that defines the point estimation $\tilde{v}$, the reduced value of the $x_{(n)}$ parameter from the $U(u,v)$ continuous uniform distribution that is derived with respect to the change point, *v* , of the OLS linear regression analysis in SPRE, as defined in equation (7). As mentioned above, the maximum spacing estimator of $\theta_0$ is a value that maximizes the logarithm of the geometric mean of the sample spacing. Therefore, statistic $S_n$ will be given here as:

$$S_n(u,v)=\frac{1}{n+1}\left\{\ln(x_{(1)}-u)+\ln(v-\tilde{x}_{(n,\tilde{v})})-(n+1)\ln(v-u)+\sum_{i=2}^{n}\ln(x_{(i)}-x_{(i-1)})\right\} \quad (8)$$

The first 3 terms that depend upon $u,v$ will be differentiated resulting in a linear system. Solving yields the maximum spacing estimate (MSE) for $u,v$ following Cheng & Amin (1983), derived to define $\tilde{x}_{(n,\tilde{v})}$ at u and v:

$$u=\frac{nx_{(1)}-\tilde{x}_{(n,\tilde{v})}}{n-1}, \text{ and } v=\frac{n\tilde{x}_{(n,\tilde{v})}-x_{(1)}}{n-1}, \text{ when } \tilde{x}_{(n,\tilde{v})} \text{ is the initial unknown parameter} \quad (9a.b)$$

*2.5 The Extended Solution for the Maximum Spacing at the Change Point of the SPRE model*

Solving for the extended solution involves 3 analytical steps.

*Step 1*: Solve for $\tilde{x}_{(n,\tilde{v})}$ including the parameters $u \text{ and } v$ from equations (9a,b):

$$\tilde{x}_{(n,\tilde{v})}=-u(n-1), \text{ and } \tilde{x}_{(n,\tilde{v})}=\frac{v(n-1)}{n}, \text{ when } \tilde{x}_{(n,\tilde{v})} \text{ is unknown} \quad (10a,b)$$

While $\tilde{x}_{(n,\tilde{v})}$ is the minimum point estimation for the minimum variance unbiased estimators (MVUE) for the continuous distribution, the $u,v$ outcomes are known at 0,0 and the change point respectively of the OLS line from the SPRE model, when 'n' is the sample number of treatment data to the change point. In this analysis, the outcome for the endpoint $u$ is known to be 1.00 at the initial data point of the OLS line, (Figure 1) where the value of $x_{(1)}$ equals 0.0 and n is equal to 1.0 at the beginning of the analysis. The equation for $\tilde{x}_{(n,\tilde{v})}=0$ is invariant at the initial data input and a trivial solution because it does not include any of the patients' responses to treatment for prostate cancer that are present throughout the OLS domain at the value of the change point. Therefore the equation derived for $\tilde{x}_{(n,\tilde{v})}$ from *v* will be used to derive the reduced value of $\tilde{v}$ as $\tilde{x}_{(n,\tilde{v})}$ which will be used to derive the maximum spacing from known outcome OLS data at the change point. In the case of biochemical freedom from failure for prostate cancer, $v=0.84$ (the survival outcome at the change point in months), n= 23 (the number of data points on the 'x' axis of the OLS line),

$\tilde{x}_{(n,\tilde{v})}$ is the outcome value at the derived point and $\hat{x}_{(1)} = 0$. For the prostate cancer analysis, $\tilde{x}_{(n,\tilde{v})}$ is determined using the boundary value of the linear analysis, in this case where $x_{(n)} = n$ = 23 months. Then, rounded to the precision of the input data:

$$\tilde{x}_{(n,\tilde{v})} = \frac{v(n-1)}{n}, \text{ and } \tilde{x}_{(n,\tilde{v})} = .80 \text{ (rounded)} \tag{11}$$

*Step 2*: To transpose $\tilde{x}_{(n,\tilde{v})}$ to the 'x' axis results in the point estimation of the number of months that yield the lower limit of the minimum-variance unbiased estimator (MVUE) maximum spacing on the 'x' axis. The objective is to determine the reduced value from the change point, given here as $\hat{x}_{(n)}$ initially using the uniform distribution $U(u,v)$, calculated from $\tilde{x}_{(n,\tilde{v})}$ and also denoted on the 'x' axis:

$$\hat{x}_{(n)} = \frac{1}{v}\left(n \cdot \tilde{x}_{(n,\tilde{v})}\right) = 21.9 \tag{12}$$

*Step 3*: The final value of the spacing $\hat{x}_{(sp)}$ is given by the distance between $\hat{x}_{(n)}$ and the number of sample treatments to the change point where $x_{(n)} = n$ on the 'x' axis. Then:

$$\hat{x}_{(sp)} = x_{(n)} - \hat{x}_{(n)}, \Rightarrow 23 - 21.9 \tag{13}$$

And:

$$\hat{x}_{(sp)} = 1.1 \tag{14}$$

In this extended maximum spacing method, $n$ is known at the SPRE change point. From equation (14), the maximum spacing forward of the change point in the SPRE model = 1.1 on a finite number line. From equation (10a), $\tilde{x}_{(n,\tilde{v})}$ for $u$ will always equal 0.0, and $\tilde{x}_{(n,\tilde{v})}$ for $v$, given in Table 1, will yield rounded identical maximum spacing results of $\hat{x}_{(sp)}$ = 1.00 at lower values of n = 20 – 4. The overall result is important when the SPRE model is used to determine the maximum spacing in personalized medicine for values of the change point much lower than $x_{(n)} = n = 23$.

*2.6 Simulations of the Continuous Uniform Distribution for 'n' at Varying Change Points*

Since in this analysis $n$ is a small number of months to the change point, a simulation has been conducted of the continuous uniform distribution in R with runif(n, min, max) when the default values are min = 0 and max = 1. These values are the lower and upper limits of the distribution, which must be finite. In this analysis, n = 23, the number of months of survival at the change point. Here, the mean = ½ for all the simulations, where the distribution of averages of 23 uniform distributions is investigated with 1000 simulations.

It can be seen that the theoretical mean is centered at 0.5, and the distribution of these sample means is centered at 0.4792. Where the distribution of averages of 10 uniform distributions is investigated with 1000 simulations, the distribution of the sample means = 0.4978, and for the distribution of averages of 5 uniform distributions the sample

means = 0.5227. The comparative differences between these simulations of means and the theoretical means are 0.0416, 0.0044, and 0.0434 respectively. These differences are relatively small even at the very smallest value of n = 5, and negligible at n = 10. Therefore, the distribution of averages of uniform distributions may be considered to be stable for these values at the change point of the OLS linear regression, and the relatively small comparative differences indicate that the OLS can be modeled as the continuous uniform distribution $U(u,v)$. The results are given in Table 1 and the graph in Figure 2. The R Core Team (2013) for these simulations can be seen in Appendix A.

Table 1. Maximum spacing $\hat{x}_{(sp)}$ for a summary of 'n-change point' input time values in months

| **n** | $\tilde{x}_{(n,\tilde{v})} = \frac{v(n-1)}{n}$ | $\hat{x}_{(n)} = \frac{1}{v}\left(n \cdot \tilde{x}_{(n,\tilde{v})}\right)$ | $\hat{x}_{(sp)} = x_{(n)} - \hat{x}_{(n)}$ | $\hat{x}_{(sp)}$ |
|---|---|---|---|---|
| **23** | 0.803 use 0.80 | 21.9 | 23 – 21.9 | 1.1 |
| **20** | 0.798 | 19.0 | 20-19 | 1 |
| **16** | 0.825 | 15.0 | 16-15 | 1 |
| **12** | 0.843 | 10.99 use 11.0 | 12-11 | 1 |
| **8** | 0.849 | 7.0 | 8-7 | 1 |
| **4** | 0.746 | 2.99 use 3.0 | 4-3 | 1 |

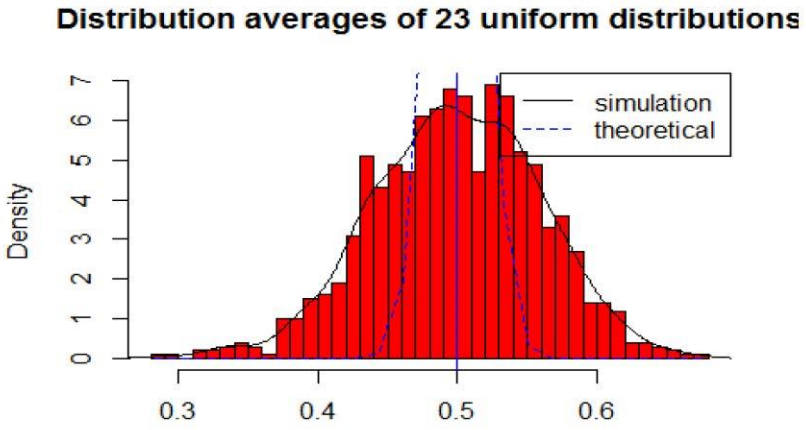


Figure 2. The Comparative Simulations of the Continuous Uniform Distribution

2.6.1 Interpretation from the Simulations for 'n'at Varying Change Points

The summary for $\hat{x}_{(sp)}$ clearly shows that at the change point for long-term change point values in prostate cancer, the maximum spacing is 1.1. Furthermore, the results of change points from 4 – 20 in value (in time or session numbers) support earlier and ongoing pilot study data having change points from 4 - 14 using a maximum spacing of 1.00. The values in Table 2 point to the use of varying maximum spacing for early, mid-term and very long-term values of the change point.

## 3. Results

### *3.1 An Important Consideration*

It should be noted that the maximum spacing (MSE) is invariant with respect to the value of time on the OLS line because, as noted earlier, all intervals of the same length on the distribution's support are equally probable. It should also be noted that the shape and scale parameters of the Weibull ratio are estimated from the OLS at the change point where the point estimates begin. Therefore, the maximum spacing will be used for all time values extending to the point estimates from $time = \tau$ at the change point to the asymptote for point estimates for prostate cancer. In Table 2, every $5^{th}$ point estimate is listed to summarize the estimated outcomes.

Table 2. Point estimates from the change point to 89 months near the end of the study for every 5th estimate

| t | $\tau$ | $R$ | $\hat{\theta}_t$ , outcomes |
|---|---|---|---|
| **23** | 23 | .632 | **.84 (from data)** |
| **27.4** | 23 | .9942 | **.807** |
| **31.8** | 23 | .9951 | **.791** |
| **36.2** | 23 | .9958 | **.777** |
| **40.6** | 23 | .9963 | **.765** |
| **45** | 23 | .9968 | **.755** |
| **49.4** | 23 | .9971 | **.747** |
| **53.8** | 23 | .9974 | **.739** |
| **58.2** | 23 | .9976 | **.731** |
| **62.6** | 23 | .9978 | **.723** |
| **67** | 23 | .998 | **.718** |
| **71.4** | 23 | .9981 | **.714** |
| **75.8** | 23 | .9982 | **.710** |
| **80.2** | 23 | .9984 | **.706** |
| **84.6** | 23 | .9985 | **.702** |
| **89** | **23** | **.9986** | **.698** |

*3.2 Related values of Predictions to Test Data*

The relative error from the final point estimate given in Table 1 to the data = 1.15%. This is an excellent result for point estimates given from the change point at 23 months to 89 months. The results are shown in Figure (3).

## 4. Discussion

"Modern transperineal brachytherapy techniques, using ultrasound guidance with either iodine-125 or palladium-103, offer good alternatives for the management of localized prostate cancer in appropriately selected patients," quoted from Potters (2003). "For those men for whom watchful waiting is not an acceptable option, brachytherapy is the treatment that results in the least impairment of patient's life-style and sexual function" (Potters, 2003). As a result of these treatment methods, the use of maximum spacing derived from the extended MSE for the SPRE model becomes particularly important. This maximum spacing then defines the gaps between point estimates at each time-dependent predicted outcome from the change point forward and results in a semiparametric ratio estimator that is reliable and repeatable. This is important when the estimates are derived from a change point in short-term linear regression of the patients' data. The long-term predictions may extend out 7.42 years, when considering freedom from biochemical failure in prostate cancer given in this study. This analysis indicates that an important application of this extended maximum spacing estimator from the change point of the SPRE model is to predict the survival of prostate cancer patients in large-scale trials research, in clinical medicine or in personalized medicine.

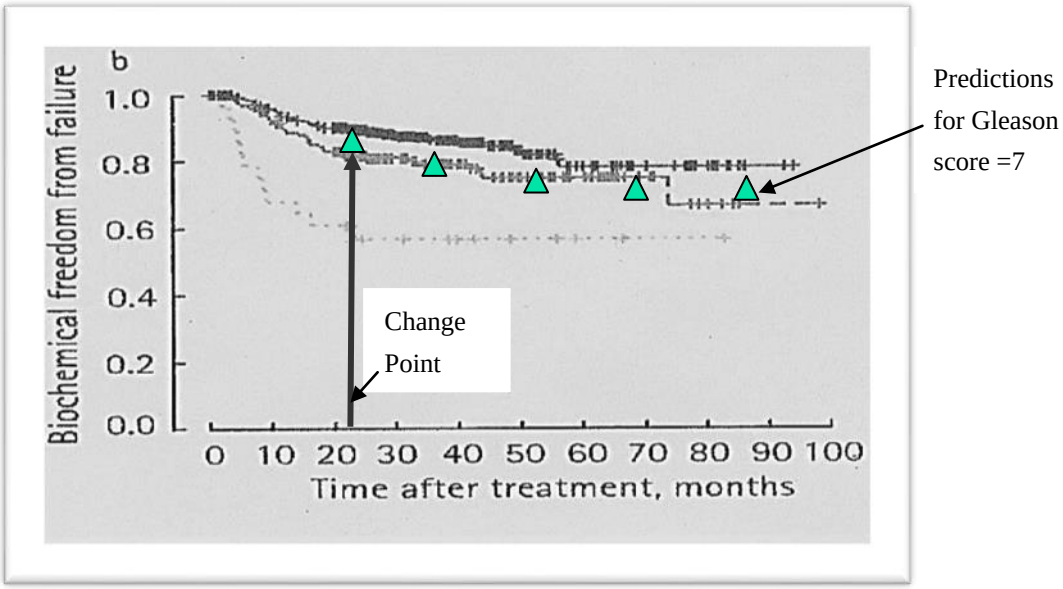


Figure 3. The results of MSE predictions from SPRE for 23 – 89 months.

(Permission granted: Joseph, et al (2004). Prostate-specific antigen relapse-free survival in patients with localized prostate cancer treated by brachytherapy. BJU International; 94) [9].)

**References**

American Cancer Society. (2009). *What are the key statistics about prostate cancer?* Retrieved October 3, 2009 at: www.cancer.org/docroot/CRI/content/CRI_2_4_1X/What_are_the_key_statistics_for_prostate_cancer_36.asp on 10/03/2009.

Cheng, R.C.H., & Amin, N.A.K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *J. R. Statist. Soc. B, 45*(3), 394-403.

Cheng, R. C. H., & Stephens, M. A. (1989). A goodness-of-fit test using Moran's statistics with estimated parameters. *Biometrika*, *76*(2), 385-92. http://dx.doi.org/10.1093/biomet/76.2.385

Eggener, S. E., & Coleman, J. A. (2008). Focal treatment of prostate cancer with vascular-targeted photodynamic therapy. *Scientific World Journal*, *8*, 963-973. http://dx.doi.org/10.1100/tsw.2008.127

Eton, D. T., & Lepore, S. J. (2002). Prostate cancer and health-related quality of life. *Psychooncology*, *11*(4), 307-326. http://dx.doi.org/10.1002/pon.572

Ghosh, K., & Jammalamadaka, S. R. (2001). A general estimation method using spacings. *Journal of Statistical Planning and Inference*, *93*, 71-82. http://dx.doi.org/10.1016/S0378-3758(00)00160-9

Joseph, J., Al-Qaisieh, B., Ash, D., Bottomley, D. & Carey, B. (2004). Prostate-specific antigen relapse-free survival in patients with localized prostate cancer treated by brachytherapy. *BJU International*, *94*, 1235-1238. http://dx.doi.org/10.1111/j.1464-410X.2004.05149.

National Cancer Institute. (2009). Prostate Cancer Homepage. Retrieved September 27, 2009 at: http://www.cancer.gov/cancertopics/types/prostate.

Pastides, H. (2001). The descriptive epidemiology of cancer. In P.C. Nasca & H. Pastides (Eds.), *Fundamentals of Cancer Epidemiology (pp. 6).* Gaithersburg, MD: Aspen Publishers.

Potters, L. (2003). Permanent prostate brachytherapy in men with clinically localized prostate cancer. *Clinical Oncology, 15*, 301-315. http://dx.doi.org/10.1016/S0936-6555(03)00152-3

Ranneby, B. (1984). The Maximum Spacing Method. An estimation related to the maximum likelihood method. *Scand J Statist, 11*, 93-112.

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Tempany, C., Straus, S., Nabuhiko, H., & Haker, S. MR. *(2008*). Guided prostate interventions. *Journal of Magnetic Resonance Imaging*, *27*(2), 356-367. http://dx.doi.org/10.1002/jmri.21259

Weissman-Miller, D. (2011). Novel maximum spacing for point estimates from a semiparametric ratio estimator: With application to survival in prostate cancer. In the Joint Statistical Meeting (JSM) Proceedings, Non-Parametric Section. Miami, FL: American Statistical Association (2011).

Weissman-Miller, D. (2013). Novel point estimation from a Semiparametric Ratio Estimator (SPRE): Long-term health outcomes from short-term linear data, with application to weight loss in obesity. *International Journal of Biostatistics*, *9*(2), 175-184. http://dx.doi.org/10.1515/ijb-2012-0049

Weissman-Miller, D., Shotwell, M.P. & Miller, R.J. (2012). New single-subject and small-n design in occupational therapy: Application to weight loss in obesity. *American Journal of Occupational Therapy, 66*, 455-462. http://dx.doi.org/10.5014/ajot.2012.004788

Weissman-Miller, D. and Graham, K. 2015. Novel scale development for fear of falling and falls: Analysed using a Semiparametric Ratio Estimator (SPRE). *International Journal of Statistics and Probability*, *4*(3). http://dx.doi.org/10.5539/ijsp.v4n3p161

Wong, T. S. T., & Li, W. K. (2006). A note of the estimation of extreme value distributions using maximum product of spacings. IMS Lecture Notes Monograph Series. *Time Series and Related Topics*, *52*, 272-283. http://dx.doi.org/10.1214/074921706000001102

**Appendix A**

**R Code for Simulations at Varying Change Points**

```
---
title: "Simulation of continuous uniform distributions"
author: "Dr. D. Weissman-Miller"
date: "Sunday, February 01, 2015"
output: word_document
---
## Distribution of averages of 23 uniform distributions
```{r,echo=FALSE}
set.seed(3)
        num_sim <- 1000
         sample_size <- 23
         sim <- matrix(runif(num_sim*sample_size), num_sim, sample_size)
         rMeans <- rowMeans(sim)
         ## Plot histogram of averages
                hist(rMeans, breaks=50, prob=TRUE,
                main="Distribution averages of 23 uniform distributions", xlab="", col="red")
                lines(density(rMeans))
                abline(v= 0.5, col="blue")
                        xfit <- seq(min(rMeans), max(rMeans), length=23)
                                print(mean(xfit))
                                yfit <- dnorm(xfit, mean=1/2, sd=1/12/sqrt(sample_size))
                            lines(xfit, yfit, pch=22, col="blue", lty=2)
                            legend('topright', c("simulation", "theoretical"), lty= c(1,2),
                                col= c("black", "blue"))
```
```