LISBOA SCHOOL OF ECONOMICS & MANAGEMENT

MASTER IN ACTUARIAL SCIENCE

**Risk Models**

**30/01/2017**

Time allowed: 3 hours

**Instructions**:

1. This paper contains 8 questions and comprises 4 pages including the title page.
2. Enter all requested details on the cover sheet.
3. You have 10 minutes reading time. You must not start writing your answers until instructed to do so.
4. Number the pages of the paper where you are going to write your answers.
5. Attempt all 8 questions.
6. Begin your answer to each of the 8 questions on a new page.
7. Marks are shown in brackets. Total marks: 200.
8. Show calculations where appropriate.
9. An approved calculator may be used.
10. The distributed formulary and the Formulae and Tables for Actuarial Examinations (the 2002 edition) may be used. Note that the parametrization used for the different distributions is that of the distributed formulary.

1. For observation $i$ of a survival study you are given:

- $d_i$ is the left truncation point
- $x_i$ is the observed value if not right censored
- $u_i$ is the observed value if right censored

| Obs. ( $i$ ) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $d_i$ | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 3 | 6 | 7 |
| $x_i$ | 6 | 12 | 12 | - | - | 6 | 5 | 9 | 16 | - |
| $u_i$ | - | - | - | 5 | 4 | - | - | - | - | 9 |

a) **[10]** Determine the Kaplan-Meier (product-limit) estimate for $S(10)$ and obtain a log-transformed confidence interval (95%) for $S(10)$

b) **[10]** Estimate ${}_2q_8$ and also estimate the conditional variance of the used estimator. Explain briefly why you can't use an unconditional variance.

c) **[5]** Using Nelson-Aalen approach give another estimate for $S(10)$

d) **[15]** Assuming that the time of survival, $X$, follows an exponential distribution, obtain a maximum likelihood estimate for $\theta$ and determine a 95% confidence interval for $\theta$ based on the asymptotic distribution of the maximum likelihood estimators.

e) **[15]** Using the results obtained in the previous question and the delta method, present a 95% confidence interval for $S(10)$. Note: If you were unable to solve the previous question, assume – and clearly state it – that $\hat{\theta}=10$ and that the interval is (3.3327; 16.6673) - which is not the correct answer to the previous question.

2. **[10]** A mortality study covers $n$ lives and you know that none were censored. Let $t_k$ be the time of the $k^{th}$ death. You also know that two deaths were observed at time $t_1$ and that the same happens at time $t_2$. From other times, no two deaths occurred at the same time. Finally the Nelson-Aalen estimate of the cumulative hazard rate function at time $t_2$ was computed and we obtained $\hat{H}(t_2)=\dfrac{96}{575}$. Determine the Kaplan-Meier estimate of the survival function at time $t_7$.

3. **[15]** You observed the sample (1 1 2 2 3 3 3 3 3 3) from a population with density function $f(x)$. Using a uniform kernel with bandwidth 1 obtain an estimate for $f(2.5)$ and for $F(2.5)$.

4. **[15]** Let $(25, 110, 235, 350, 1820)$ be an observed (random) sample from a Pareto population with parameters $\alpha$ and $\theta$, both unknown. Using the method of moments, estimate both parameters. Also obtain an estimate for $E(X \wedge 500)$. Note: if you are unable to get the moment estimates for $\alpha$ and $\theta$ assume – and clearly state it – that $\tilde{\alpha} = 5$ and $\tilde{\theta} = 2000$ (which is not the correct answer) to estimate $E(X \wedge 500)$.

5. You are given data from 6 independent policies. Each policy follows the same ground up loss distribution $f(x|\alpha,\theta)$, $x > 0$, $\alpha, \theta > 0$, but the insurance contracts are not the same. The information for each policy is as follows:

| Policy | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Payment | 10 | 5 | 7 | 8 | 10 | 3 |
| Deductible | - | 1 | - | 2 | - | 3 |
| Payment limit | 10 | 5 | 10 | 20 | 20 | 10 |

As it is usually the case, losses below the deductible are not reported and losses are paid in excess of the deductible.

   a. **[10]** Write the log-likelihood function needed to get the maximum likelihood estimates of $\alpha$ and $\theta$ using the density and the survival functions of the **ground-up losses**.

   b. Using R, you got the following output (remember that the R function mln minimizes the function that is given and that the function solve compute the inverse of a matrix).

```
>minusloglik =function(param,z){
>   alpha=param[1]; theta=param[2]
    ….
>   return(-sum(ll))
>   }
> out=nlm(minusloglik,param.start,hessian=T,z=dados)
There were 12 warnings (use warnings() to see them)
> out
$minimum
[1] 13.91427
$estimate
[1]  3.92998 25.53616
$gradient
[1] -5.993540e-07  3.763327e-08
$hessian
         [,1]        [,2]
[1,]  0.43014410 -0.02902629
[2,] -0.02902629  0.09467377
```

```
$code
[1] 1
$iterations
[1] 26
> solve(out$hessian)
         [,1]       [,2]
[1,] 2.3739163  0.7278255
[2,] 0.7278255 10.7857333
```

i. **[10]** Obtain 95% confidence interval for $\theta$.
ii. **[15]** Obtain a 95% confidence interval for $\beta = \alpha\theta$.

6. The annual number of claims on a given policy has a geometric distribution with parameter $\beta$, i.e. $f(x \mid \beta) = \dfrac{\beta^x}{(1+\beta)^{x+1}}$, $x = 0,1,2,\ldots$, $\beta > 0$ .The prior density for $\beta$ is given by $\pi(\beta) = \dfrac{\alpha}{(\beta+1)^{\alpha+1}}$, $\beta > 0$ and $\alpha$ is known to be greater than 2. A sample of **size 1**, with $x_1 = 1$, was observed.
   a. **[10]** Show that the posterior distribution for $\beta$ is given by
   $\pi(\beta \mid \underline{x}) = \dfrac{(\alpha+2)(\alpha+1)\beta}{(1+\beta)^{\alpha+3}}$, $\beta > 0$
   b. **[10]** Obtain a Bayes estimate for $\beta$ assuming a 0-1 loss function.
   c. **[10]** Write the equation that needs to be solved to get a Bayes estimate for $\beta$ assuming an absolute loss function (you do not need to solve the equation, just write it)

7. **[15]** A random sample of 100 losses from a gamma distribution with $\alpha = 4$ gives the following statistics: $\min x_i = 57.72$, $\max x_i = 1046.64$, $\overline{x} = 351.53$, $s = 191.02$, $\sum_{i=1}^{100} \ln x_i = 570.25$. Use the likelihood ratio test to test $H_0 : \theta = 100$ against $H_1 : \theta \neq 100$ and conclude.

8. Assume that we observed a random samples of salaries for 30 actuaries of similar companies (similar size) in the same country.
   a. **[15]** Explain how to use the bootstrap technique to determine a 95% confidence interval for the expected salary of an actuary in this country.
   b. **[10]** Now assume that we can link to each actuary in our sample a measure of performance. Explain how to use the bootstrap to determine a 95% confidence interval for the correlation coefficient between performance and earned salary.

**SOLUTION**

1.

**a)** KM approach

| $y$ | 5 | 6 | 9 | 12 | 16 |
|---|---|---|---|---|---|
| $s$ | 1 | 2 | 1 | 2 | 1 |
| $r$ | 7 | 5 | 5 | 3 | 1 |

$$S_n(10) = \prod_{j:y_j \le 10} \frac{r_j - s_j}{r_j} = \frac{6}{7} \times \frac{3}{5} \times \frac{4}{5} = \frac{72}{175} = 0.41143$$

Greenwood's approximation:

$$\hat{\mathrm{var}}(S_n(10)) \approx S_n(10)^2 \sum_{j:y_j \le 10} \frac{s_j}{r_j\,(r_j - s_j)} = 0.41143^2 \times \left( \frac{1}{7\times 6} + \frac{2}{5\times 3} + \frac{1}{5\times 4} \right) = 0.035064$$

$$U = \exp\left( \frac{1.96 \times \sqrt{0.085224}}{0.41143 \times \ln(0.41143)} \right) = 0.3663$$

Then the 95% log-transformed CI is $(0.41143^{1/0.3663}; 0.41143^{0.3663})$, i.e. (0.0885; 0.7223)

**b)**

$${}_2q_8 = P(X \le 10 \mid X > 8) = \frac{S(8) - S(10)}{S(8)}$$

$$S_n(8) = \prod_{j:y_j \le 8} \frac{r_j - s_j}{r_j} = \frac{6}{7} \times \frac{3}{5} = \frac{18}{35} = 0.51429$$

$${}_2\hat{q}_8 = \frac{S_n(8) - S_n(10)}{S_n(8)} = \frac{0.51429 - 0.41143}{0.51429} = 0.2$$

$$\hat{\mathrm{var}}\left( {}_2\hat{q}_8 \mid S(8) = S_n(8) \right) = \left( \frac{S_n(10)}{S_n(8)} \right)^2 \sum_{j:8 < y_j \le 10} \frac{s_j}{r_j \times \left( r_j - s_j \right)} = \left( \frac{0.41143}{0.51429} \right)^2 \times \left( \frac{1}{5\times 4} \right) = 0.032$$

The variance is conditional because the unconditional variance does not exist as the estimator of $S(8)$, $S_n(8)$, can assume value 0 with a positive probability.

**c)**

$$\hat{H}(10) = \sum_{j:y_j \le 10} \frac{s_j}{r_j} = \frac{1}{7} + \frac{2}{5} + \frac{1}{5} = 0.74286$$

Then $\hat{S}(10) = e^{-\hat{H}(10)} = e^{-0.74286} = 0.47575$

**d)**

$$\ell(\theta) = \ln\left( f(6 \mid \theta) \right) + 2 \times \ln\left( f(12 \mid \theta) \right) + \ln\left( S(5 \mid \theta) \right) + \ln\left( S(4 \mid \theta) \right) +$$
$$+ \ln\left( \frac{f(6 \mid \theta)}{S(2 \mid \theta)} \right) + \ln\left( \frac{f(5 \mid \theta)}{S(3 \mid \theta)} \right) + \ln\left( \frac{f(9 \mid \theta)}{S(3 \mid \theta)} \right) + \ln\left( \frac{f(16 \mid \theta)}{S(6 \mid \theta)} \right) + \ln\left( \frac{S(9 \mid \theta)}{S(7 \mid \theta)} \right)$$

As $\ln(f(x \mid \theta)) = \ln\left( \theta^{-1} e^{-x/\theta} \right) = -\ln\theta - x/\theta$ and $\ln(S(x \mid \theta)) = \ln\left( e^{-x/\theta} \right) = -x/\theta$ we get

$$\ell(\theta) = -7\ln\theta - \frac{6}{\theta} - 2\times\frac{12}{\theta} - \frac{5}{\theta} - \frac{4}{\theta} - \frac{6}{\theta} + \frac{2}{\theta} - \frac{5}{\theta} + \frac{3}{\theta} - \frac{9}{\theta} + \frac{3}{\theta} - \frac{16}{\theta} + \frac{6}{\theta} - \frac{9}{\theta} + \frac{7}{\theta} = -7\ln\theta - \frac{63}{\theta}$$

$$\ell'(\theta) = -\frac{7}{\theta} + \frac{63}{\theta^2} \quad \text{and} \quad \ell''(\theta) = \frac{7}{\theta^2} - \frac{126}{\theta^3}$$

$$\ell'(\theta) = 0 \Leftrightarrow \frac{7}{\theta} = \frac{63}{\theta^2} \Leftrightarrow \theta = \frac{63}{7} = 9 \text{ and } \hat{\theta} = 9 \text{ as } \ell''(\hat{\theta}) = \frac{7}{81} - \frac{14}{81} = -\frac{7}{81} < 0$$

$$\hat{\text{var}}(\hat{\theta}) = -1/\ell''(\hat{\theta}) = 81/7$$

Then the 95% CI for $\theta$ is $9 \pm 1.96 \times \sqrt{81/7}$, i.e. (2.3327; 15.6673)

**e)**

$\hat{S}(10) = e^{-10/\hat{\theta}} = 0.3292$ and

$$\hat{\text{var}}(\hat{S}(10)) = \left(\frac{10}{\hat{\theta}^2} e^{-10/\hat{\theta}}\right)^2 \hat{\text{var}}(\hat{\theta}) = \left(\frac{10}{81} e^{-10/9}\right)^2 \frac{81}{7} = \frac{100\, e^{-20/9}}{81 \times 7} = 0.019113$$

Then the 95% CI for $S(10)$ is $0.3292 \pm 1.96 \times \sqrt{0.019113}$, i.e. (0.0582; 0.6002)

2.

$$\hat{H}(t_2) = \frac{96}{575} = \frac{s_1}{r_1} + \frac{s_2}{r_2} = \frac{2}{n} + \frac{2}{n-2} = \frac{2n - 4 + 2n}{n(n-2)} = \frac{4n-4}{n(n-2)}$$

Solving in order to $n$

$$\frac{96}{575} = \frac{4n-4}{n(n-2)} \Leftrightarrow 96n^2 - 192n = 2300n - 2300 \Leftrightarrow 96n^2 - 2492n + 2300 = 0$$

$$n = \frac{2492 \pm \sqrt{2492^2 - 4 \times 96 \times 2300}}{2 \times 96} = \frac{2492 \pm 2308}{2 \times 96}$$, i.e. $n = 25$ or $n = 0.9583$

As the value $n = 0.9583$ is not admissible we get $n = 25$

Then

$$S_{25}(t_7) = \frac{23}{25} \times \frac{21}{23} \times \frac{20}{21} \times \frac{19}{20} \times \frac{18}{19} \times \frac{17}{18} \times \frac{16}{17} = \frac{16}{25} = 0.64$$

3.

| $y_j$ | 1 | 2 | 3 |
|---|---|---|---|
| $p(y_j)$ | 0.2 | 0.2 | 0.6 |

$$k_y(x) = \begin{cases} 0.5 & y-1 < x < y+1 \\ 0 & \text{elsewhere} \end{cases}$$

$$K_y(x) = \begin{cases} 0 & x < y-1 \\ \int_{y-1}^{x} (1/2)\, du & y-1 \le x < y+1 \\ 1 & x \ge y+1 \end{cases} = \begin{cases} 0 & x < y-1 \\ \dfrac{x-y+1}{2} & y-1 \le x < y+1 \\ 1 & x \ge y+1 \end{cases}$$

$$\hat{f}(2.5) = \sum_{j=1}^{k} p(y_j)\, k_{y_j}(x) = 0.2 \times 0 + 0.2 \times 0.5 + 0.6 \times 0.5 = 0.4$$

$$\hat{F}(2.5) = \sum_{j=1}^{k} p(y_j)\, K_{y_j}(x) = 0.2 \times 1 + 0.2 \times \frac{2.5 - 2 + 1}{2} + 0.6 \times \frac{2.5 - 3 + 1}{2} = 0.2 + 0.2 \times 0.75 + 0.6 \times 0.25 = 0.56$$

4. $E(X \mid \alpha,\theta)=\dfrac{\theta}{\alpha-1}$; $E(X^2 \mid \alpha,\theta)=\dfrac{2\theta^2}{(\alpha-1)(\alpha-2)}$ and $\bar{x}=508$; $t=\dfrac{\sum_{i=1}^{n} x_i^2}{n}=700570$

To obtain our estimates we need to solve

$$\begin{cases} \dfrac{\theta}{(\alpha-1)}=\bar{x} \\ \dfrac{2\theta^2}{(\alpha-1)(\alpha-2)}=t \end{cases} \Leftrightarrow \begin{cases} \theta=\bar{x}(\alpha-1) \\ 2\bar{x}^2(\alpha-1)^2=t(\alpha-1)(\alpha-2) \end{cases} \Leftrightarrow \begin{cases} \theta=\bar{x}(\alpha-1) \\ 2\bar{x}^2(\alpha-1)=t(\alpha-2) \end{cases} \Leftrightarrow \begin{cases} \theta=\bar{x}(\alpha-1) \\ \alpha=\dfrac{2t-2\bar{x}^2}{t-2\bar{x}^2} \end{cases}$$

Then the estimates are $\tilde{\alpha}=4.798321$ and $\tilde{\theta}=1929.547$.

Consequently, as $E(X \wedge 500)=\dfrac{\theta}{\alpha-1}\left(1-\left(\dfrac{\theta}{500+\theta}\right)^{\alpha-1}\right)$, we get

$$\hat{E}(X \wedge 500)=\frac{\hat{\theta}}{\hat{\alpha}-1}\left(1-\left(\frac{\hat{\theta}}{500+\hat{\theta}}\right)^{\hat{\alpha}-1}\right)=508\times(1-0.7942^{3.798321})=296.278$$

5.

**a.**

$$\ell(\alpha,\theta)=\ln S(10 \mid \alpha,\theta)+\ln S(6 \mid \alpha,\theta)-\ln S(1 \mid \theta)+\ln f(7 \mid \alpha,\theta)+\ln f(10 \mid \alpha,\theta)-\ln S(2 \mid \alpha,\theta)+$$
$$+\ln f(10 \mid \alpha,\theta)+\ln f(6 \mid \alpha,\theta)-\ln S(3 \mid \alpha,\theta)$$

**b.**

i. from the output $\hat{\theta}=25.5362$ and $\hat{\mathrm{var}}(\hat{\theta})=10.7857$ and then the 95% confidence interval is given by $25.5362 \pm 1.96\sqrt{10.7857}$, i.e. (19.099; 31.973)

ii.

$\beta=\alpha\theta$

$\hat{\beta}=\hat{\alpha}\hat{\theta}=3.9300\times 25.5362=100.3566$

Using the delta method

$$\mathrm{var}(\hat{\beta})=\begin{bmatrix}\hat{\theta} & \hat{\alpha}\end{bmatrix}\begin{bmatrix}\hat{\mathrm{var}}(\hat{\alpha},\hat{\theta})\end{bmatrix}\begin{bmatrix}\hat{\theta} \\ \hat{\alpha}\end{bmatrix}=\begin{bmatrix}25.5362 & 3.9300\end{bmatrix}\begin{bmatrix}2.3739 & 0.7278 \\ 0.7278 & 10.7857\end{bmatrix}\begin{bmatrix}25.5362 \\ 3.9300\end{bmatrix}$$
$$=\begin{bmatrix}63.48064 & 60.97305\end{bmatrix}\begin{bmatrix}25.5362 \\ 3.9300\end{bmatrix}=1860.678$$

And the 95% confidence interval is $100.3566 \pm 1.96\times\sqrt{1860.678}$, i.e (15.8109; 184.9023)

6. a)

$$L(\beta)=\beta^{x_1}(1+\beta)^{-x_1-1}$$

$$\pi(\beta)=\frac{\alpha}{(\beta+1)^{\alpha+1}}\propto(\beta+1)^{-\alpha-1}$$

$$\pi(\beta\,|\,\underline{x})\propto(\beta+1)^{-\alpha-1}\,\beta^{x_1}\,(1+\beta)^{-x_1-1}=\beta^{x_1}\,(1+\beta)^{-x_1-\alpha-2}\ ,\ \beta>0$$

As (integrating by parts)

$$\int_0^\infty\beta(1+\beta)^{-\alpha-3}\,d\beta=\left(-\frac{\beta(1+\beta)^{-\alpha-2}}{\alpha+2}\right]_0^\infty+\frac{1}{\alpha+2}\int_0^\infty(1+\beta)^{-\alpha-2}\,d\beta$$

$$=\frac{1}{\alpha+2}\left(-\frac{(1+\beta)^{-\alpha-1}}{\alpha+1}\right]_0^\infty=\frac{1}{\alpha+2}\times\frac{1}{\alpha+1}$$

We get the posterior

$$\pi(\beta\,|\,\underline{x})=\frac{(\alpha+2)(\alpha+1)\beta}{(1+\beta)^{\alpha+3}}\ ,\ \beta>0$$

**b)** The Bayes estimate is the mode of the posterior distribution.

As $\pi(\beta\,|\,\underline{x})=\frac{(\alpha+2)(\alpha+1)\beta}{(1+\beta)^{\alpha+3}}$ , $\beta>0$

Let us maximize the log of the posterior

$$h(\beta)=\ln\left(\frac{(\alpha+2)(\alpha+1)\beta}{(1+\beta)^{\alpha+3}}\right)=\ln(\alpha+2)+\ln(\alpha+1)+\ln\beta-(\alpha+3)\ln(1+\beta)$$

$h'(\beta)=\frac{1}{\beta}-\frac{(\alpha+3)}{1+\beta}$ and $h'(\beta)=0\Leftrightarrow\frac{1}{\beta}=\frac{(\alpha+3)}{1+\beta}\Leftrightarrow\frac{1}{\beta}+1=\alpha+3\Leftrightarrow\beta=\frac{1}{\alpha+2}$

We can check that this is a minimum as

$h''(\beta)=-\frac{1}{\beta^2}+\frac{(\alpha+3)}{(1+\beta)^2}$ and then

$$h''(\frac{1}{\alpha+2})=-(\alpha+2)^2+\frac{(\alpha+3)(\alpha+2)^2}{(\alpha+3)^2}=(\alpha+2)^2\left(-1+\frac{1}{(\alpha+3)}\right)<0$$

The Bayes estimate against a 0-1 loss function is then $\hat{\beta}_{0-1}=\frac{1}{\alpha+2}$

c.

The Bayes estimate is the median of the posterior distribution.

As $\pi(\beta\,|\,\underline{x})=\frac{(\alpha+2)(\alpha+1)\beta}{(1+\beta)^{\alpha+3}}$ , $\beta>0$, we need to find $\hat{\beta}_{Abs}$ such that

$$\int_0^{\hat{\beta}_{Abs}}\pi(\beta\,|\,\underline{x})\,d\beta=\frac{1}{2}\Leftrightarrow\int_0^{\hat{\beta}_{Abs}}\frac{(\alpha+2)(\alpha+1)\beta}{(1+\beta)^{\alpha+3}}\,d\beta=0.5$$

7. $f(x\,|\,\theta)=\frac{x^3\,e^{-x/\theta}}{6\theta^4}$ , $x>0$ , $\theta>0$

$H_0:\theta=100$ $\qquad\qquad H_1:\theta\neq100$

$$\ell(\theta)=\sum_{i=1}^{100}\left(-\ln6-4\ln\theta+3\ln x_i-\frac{x_i}{\theta}\right)=-100\ln6-400\ln\theta+3\sum_{i=1}^{100}\ln x_i-\frac{\sum_{i=1}^{100}x_i}{\theta}$$

$$=-179.176-400\ln\theta+1710.754-35153/\theta=1531.578-400\ln\theta-35153/\theta$$

$$\ell_0=\ell(100)=1531.578-400\ln100-35153/100=-662.02$$

$$\ell'(\theta) = -\frac{400}{\theta} + \frac{35153}{\theta^2} \qquad \ell'(\theta) = 0 \Leftrightarrow \frac{400}{\theta} = \frac{35153}{\theta^2} \Leftrightarrow \theta = \frac{35153}{400} = 87.8825$$

$$\ell_1 = \ell(87.8825) = 1531.578 - 400 \ln 87.8825 - 35153 / 87.8825 = -658.82$$

The observed value for the test statistic is $T_{obs} = 2 \times (-658.82 + 662.02) = 6.4$

The critical value is $q_\alpha = 3.8415$ (quantile 0.95 of a chi-square distribution with 1 degree of freedom).

Then we reject $H_0 : \theta = 100$.

8.

a.

One possible method

i. Define $NR$, the number of replicas to be used (large value)

ii. For each replica $j = 1, 2, \cdots, NR$

   1. Resample from the original sample with replacement. One possible method is to define 30 uniforms, $u_1, u_2, \cdots, u_{30}$ and, for each $u_i$ pick a value $y_i$ using: $y_i = x_k$ if $\frac{k-1}{30} < u_i \le \frac{k}{30}$
   2. Compute the pseudo sample average using $\bar{x}_j^{(B)} = \frac{\sum_{i=1}^{30} x_i}{30}$ and keep it.

iii. Pick the appropriate quantiles (0.025 and 0.975) from the set of values $\bar{x}_j^{(B)}$, $j = 1, 2, \cdots, NR$, and define the confidence interval.

b.

Now each observation is a pair of values $(x_i, y_i)$ where $x_i$ is the salary and $y_i$ is the performance. Using the same approach as in the previous question,

i. Define $NR$, the number of replicas to be used (large value)
ii. For each replica $j = 1, 2, \cdots, NR$
    1. Resample from the original sample with replacement. One possible method is to define 30 uniforms, $u_1, u_2, \cdots, u_{30}$ and, for each $u_i$ pick the pair $z_i = (x_i, y_i)$ using: $z_i = (x_k, y_k)$ if $\frac{k-1}{30} < u_i \le \frac{k}{30}$
    2. Now compute the correlation coefficient between $x$ and $y$ and keep it as $r_j^{(B)}$.
iii. Pick the appropriate quantiles (0.025 and 0.975) from the set of values $r_j^{(B)}$, $j = 1, 2, \cdots, NR$, and define the confidence interval.