Dr. Quapp: Statistik für Mathematiker mit SPSS

Hinweise zur 15. Übung No 2 + 3 - Exponentialverteilung -

2.] a) Es sei $U \sim U(0,1)$. Zeigen Sie, daß $Y := -ln(1-U)$ und $Y := -ln(U)$ exponentialverteilt sind mit dem Paramter 1.
b) Erzeugen Sie mit dem Resultat aus a) 111 Zufallszahlen in SPSS.
c) Erzeugen Sie mit dem in SPSS programmierten Zufallsgenerator 111 exponentialverteilte Zufallszahlen. Vergleichen Sie die Resultate; insbesondere Modalwert und Median.

Zu a) Es sei $y = h(u) = -ln(u)$, also ist $e^y = e^{ln(\frac{1}{u})} = \frac{1}{u}$ und $u = h^{-1}(y) = e^{-y}$. Die Ableitung ist $h' = -1/u$. Somit wird mit der Dichte 1 von u über dem Intervall von (0,1)

$$f_Y(y) = f_U(h^{-1}(y))\frac{1}{|h'(h^{-1}(y))|} = 1_I \bullet \frac{1}{|\frac{1}{e^{-y}}|} = e^{-y}$$

also die Dichte der Exponentialverteilung mit dem parameter λ=1:

$$f(y) = \begin{cases} \lambda e^{-\lambda y} & \text{f''ur } 0 \le y \\ 0 & \text{sonst} \end{cases}$$

Eine analoge Rechnung gilt für die zweite Version $Y = -ln(1-U)$.

zu b+c) Mit u= RV.UNIFORM(0,1) und der Berechnung der zwei Versionen von Y1 und Y2 kann man davon ausgehen, daß man zufallsverteilte Zahlen erhalten hat, die einer Exponantialverteilung folgen. Dies beruht auf der Inversionsmethode. Verschafft man sich noch mit ee=RV.EXP(1) entsprechende zufallsverteilte Zahlen von SPSS, so kann man alle drei vergleichen:
In einem Test ergab sich als Resultat:
Für Y1 der Mittelwert=0.961, Median=0.748, und Modalwert=0.004 .
Für Y2 der Mittelwert=1.151, Median=0.641, und Modalwert=0.009 .
Für EE der Mittelwert=0.950, Median=0.611, und Modalwert=0.010 .
Das ist natürlich nur ein Vergleich der Zufallszahlen, noch dazu mit einer "geringen" Anzahl, aber die Übereinstimmung ist schon gut. Mittels Darstellung durch Histogramme kann man sich auch graphisch davon überzeugen.
Wenn die Exponentialverteilung zur Beschreibung des radioaktiven Zerfalls verwendet wird, entspricht der Median der Halbwertszeit, der Zeit, in der die Hälfte einer Menge eines radioaktiven Elementes umgewandelt worden ist. Es ist mit der Verteilungsfunktion

$$F(t) = 1 - e^{-\lambda t} \text{ sofort } F(t_{1/2}) = 1/2, \text{ also } t_{1/2} = \frac{1}{\lambda} ln(2) \ .$$

Der Erwartungswert, also die durchschnittliche "Lebenserwartung" eines Elementes, ist

$$\int_0^\infty t\,\lambda\, e^{-\lambda t}\, dt = \frac{1}{\lambda}$$

3.] Es seien $X_i \sim Exp(\lambda)$, $i = 1, 2, ..., n$.
a) Zeigen Sie, daß $\bar{X}$ eine erwartungstreue Schätzung für $1/\lambda$ ist.
b) Zur Schätzung von $P(X_i < 1)$ wählt man die relative Häufigkeit der entsprechenden Ereignisse. Welche Eigenschaften hat diese Schätzung?
c) Zur Schätzung von $P(X_i < 1)$ gibt es weiterhin den Ansatz

$$1 - e^{-1/\bar{X}} .$$

Welche Schätzung ist vorzuziehen?
d) Überprufen Sie mit SPSS die Aussagen aus a), b) und c).

zu a) Sind n Zufallszahlen X_i wie oben gegeben, so ist der Mittelwert $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ und sein Erwartungswert ist nach dem schwachen Gesetz der großen Zahlen

$$E\,\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} E\,X_i = \frac{n}{n} E\,X_1 = \frac{1}{\lambda} .$$

zu b) Die Exponantialverteilung beschreibt Ausfallprozesse, d.h. man wird besonders an der "vorderen" Kante der Verteilung interessiert sein: Wie viele Geräte fallen nach dem ersten Einschalten etwa für $x < 1$ aus?
Sei also

$$\delta = P(X_j < 1) = \int_0^1 \lambda e^{-\lambda x} dx = 1 - e^{-\lambda}$$

dann ist unsere Schätzung mittels der relativen Häufigkeit

$$\hat{\delta} = \#\{j : X_j < 1\}\,\frac{1}{n} = \sum_{j=1}^{n} I_{\{X_j<1\}}\,\frac{1}{n} =: I_n ,$$

und dessen Erwartungswert ist sozusagen mit einem Bernoulli-Experiment bestimmbar: Wenn n Versuche durchgeführt werden, wobei jeweils $p = 1 - e^{-\lambda}$ ist, und entsprechend $q = e^{-\lambda}$, so ist

$$E\,I_n = \sum_{k=1}^{n} \frac{k}{n}\binom{n}{k}(1 - e^{-\lambda})^k (e^{-\lambda})^{n-k}$$

$$= \sum_{r=0}^{m+1}\binom{m}{r}(1 - e^{-\lambda})^{r+1}(e^{-\lambda})^{m-r} = (1 - e^{-\lambda})\sum_{r=0}^{m+1}\binom{m}{r}(p)^{r+1}(q)^{m-r} = (1 - e^{-\lambda}) ,$$

mit $m = n - 1$ und $r = k - 1$ und wegen $\frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}$ und $\binom{m}{m+1} = 0$. Also ist

$$E\,\hat{\delta} = E\,I_{\{X_j<1\}} = P(X_j < 1) = \delta ,$$

also ist $\hat{\delta}$ erwartungstreu.
Die Varianz ist:

$$Var(\hat{\delta}) = \frac{1}{n} Var(I_{\{X_j<1\}}) = \frac{1}{n}\,\delta\,(1 - \delta) .$$

zu c) Sei $\hat{\delta}_1 = 1 - e^{-1/\bar{X}} = 1 - e^{-n/(X_1+...+X_n)}$.
Dessen Erwartungswert ist mit unseren Mitteln unberechenbar. Also bleibt nur der "experimentelle" Test mit SPSS.