

## 作図のための統計処理言語 R のスクリプト

『気象庁震度データベースを使用した地震リテラシー教育－「地震の相場観」を鍛える－』（高知大学工学部 村上英記）中のポアソン乱数や指数乱数を使ったシミュレーションと作図のための統計処理言語 R のスクリプトを参考までに示す。統計処理言語 R(ver. 3. 6. 0)と開発環境 RStudio(ver. 1. 1. 463)を使用して実行した。また、使用したパッケージは tidyverse(ver. 1. 2. 1), cowplot(ver. 0. 94), readxl(ver. 1. 3. 1)である。

### (1) 図 4, 5, 8 の作図用スクリプト

```
1   rm(list=ls())
2   library(tidyverse)
3   library(cowplot)
4   library(readxl)
5   #----- read data
6   SIdata <- read_excel("kochiHonmachi.xlsx", sheet=1, col_names=FALSE, skip=1)
7   colnames(SIdata)<-c("Year", "SI1", "SI2", "SI3", "SI4", "SI5L", "SI5U", "SI6L", "SI6U", "SI7")
8   #SIdata: Year SI1 SI2 SI3 SI4 ... SI7
9   num<-nrow(SIdata)
10  ss1<-apply(SIdata[, 2:10], 1, sum)
11  ss2<-apply(SIdata[, 3:10], 1, sum)
12  ss3<-apply(SIdata[, 4:10], 1, sum)
13  ss4<-apply(SIdata[, 5:10], 1, sum)
14  #
15  DATA<-
    data.frame(Y=as.numeric(str_sub(SIdata$Year, start=1, end=4)), SIdata$Year, SSI1=ss1, SSI2=ss2, SSI3=ss3, SSI4=ss4)
16  #--- Draw frequency distributions of felt earthquakes (Fig. 4)
17  p1<-
    ggplot(data=DATA, aes(x=Y, y=SSI1))+geom_bar(stat="identity")+labs(x="year", y="frequency")+geom_hline(y
    intercept=median(DATA$SSI1), linetype="dashed", colour="red", size=0.8)+geom_vline(xintercept=1946, linet
    ype="dashed", colour="blue", size=0.8)+annotate("text", label=">=SI1", size=5, x=2000, y=70)+theme(axis.tit
    le.x=element_blank())
18  p2<-
    ggplot(data=DATA, aes(x=Y, y=SSI2))+geom_bar(stat="identity")+labs(x="year", y="frequency")+geom_hline(y
    intercept=median(DATA$SSI2), linetype="dashed", colour="red", size=0.8)+geom_vline(xintercept=1946, linet
    ype="dashed", colour="blue", size=0.8)+annotate("text", label=">=SI2", size=5, x=2000, y=25)+theme(axis.tit
    le.x=element_blank())
19  p3<-
    ggplot(data=DATA, aes(x=Y, y=SSI3))+geom_bar(stat="identity")+labs(x="year", y="frequency")+geom_hline(y
```

```

intercept=median(DATA$SSI3), linetype="dashed", colour="red", size=0.8)+geom_vline(xintercept=1946, linet
ype="dashed", colour="blue", size=0.8)+annotate("text", label=">=SI3", size=5, x=2000, y=3)+theme(axis.titl
e.x=element_blank())
20   p4<-
ggplot(data=DATA, aes(x=Y, y=SSI4))+geom_bar(stat="identity")+labs(x="year", y="frequency")+geom_hline(y
intercept=median(DATA$SSI4), linetype="dashed", colour="red", size=0.8)+geom_vline(xintercept=1946, linet
ype="dashed", colour="blue", size=0.8)+annotate("text", label=">=SI4", size=5, x=2000, y=3)
21   p<-cowplot::plot_grid(p1, p2, p3, p4, labels=c("a", "b", "c", "d"), align="v", nrow=4)
22   print(p)
23   #---- Draw histograms of felt earthquakes (Fig.5)
24   xmax1<-max(DATA$SSI1)
25   xmax2<-max(DATA$SSI2)
26   xmax3<-max(DATA$SSI3)
27   xmax4<-max(DATA$SSI4)
28   xmax<-max(xmax1, xmax2, xmax3, xmax4)
29   h1<-
ggplot(data=DATA, aes(x=SSI1))+geom_histogram(binwidth=1)+labs(x="events/year", y="frequency")+xlim(-
1, xmax+1)+annotate("text", label=">=SI1", size=5, x=xmax*0.5, y=10)+theme(axis.title.x=element_blank())
30   h2<-
ggplot(data=DATA, aes(x=SSI2))+geom_histogram(binwidth=1)+labs(x="events/year", y="frequency")+xlim(-
1, xmax+1)+annotate("text", label=">=SI2", size=5, x=xmax*0.5, y=20)+theme(axis.title.x=element_blank())
31   h3<-
ggplot(data=DATA, aes(x=SSI3))+geom_histogram(binwidth=1)+labs(x="events/year", y="frequency")+xlim(-
1, xmax+1)+annotate("text", label=">=SI3", size=5, x=xmax*0.5, y=35)+theme(axis.title.x=element_blank())
32   h4<-
ggplot(data=DATA, aes(x=SSI4))+geom_histogram(binwidth=1)+labs(x="events/year", y="frequency")+xlim(-
1, xmax+1)+annotate("text", label=">=SI4", size=5, x=xmax*0.5, y=60)
33   h<-cowplot::plot_grid(h1, h2, h3, h4, labels=c("a", "b", "c", "d"), align="v", nrow=4)
34   print(h)
35   #---- Calculaion of 2.5% and 97.5% points of Poisson distribution
36   lambda1<-mean(DATA$SSI1)
37   lambda2<-mean(DATA$SSI2)
38   lambda3<-mean(DATA$SSI3)
39   lambda4<-mean(DATA$SSI4)
40   var1<-var(DATA$SSI1)
41   var2<-var(DATA$SSI2)
42   var3<-var(DATA$SSI3)
43   var4<-var(DATA$SSI4)

```

```

44 #---- simulation
45 N<-10000
46 n<-num
47 #----- ss1
48 lambda<-lambda1
49 binmax<-xmax
50 y<-array(1:N*(binmax+1),dim=c(N,binmax+1))
51 for (i in 1:N){
52   x<-rpois(n, lambda)
53   u<-hist(x, breaks=c(-1:binmax), plot=FALSE)
54   y[i, 1:(binmax+1)]<-u$counts
55 }
56 i25<-round(N*0.025, digits=0)
57 i75<-round(N*0.975, digits=0)
58 D<-rep(0, N)
59 UL<-rep(0, binmax+1)
60 UM<-rep(0, binmax+1)
61 UH<-rep(0, binmax+1)
62 for (j in 1:(binmax+1)){
63   D<-sort(y[1:N, j])
64   Dave<-mean(D)
65   UL[j]<-D[i25]
66   UM[j]<-Dave
67   UH[j]<-D[i75]
68 }
69 xc<-c(0:binmax)
70 SP1<-data.frame(EventRate=xc, frequency=UL, Poisson_D="2.5%")
71 SP2<-data.frame(EventRate=xc, frequency=UM, Poisson_D="mean")
72 SP3<-data.frame(EventRate=xc, frequency=UH, Poisson_D="97.5%")
73 DB1<-rbind(SP3, SP2, SP1)
74 rm(SP1, SP2, SP3, y)
75 #----- ss2
76 lambda<-lambda2
77 binmax<-xmax
78 y<-array(1:N*(binmax+1), dim=c(N, binmax+1))
79 for (i in 1:N){
80   x<-rpois(n, lambda)
81   u<-hist(x, breaks=c(-1:binmax), plot=FALSE)

```

```

82     y[i, 1:(binmax+1)]<-u$counts
83 }
84 i25<-round(N*0.025,digits=0)
85 i75<-round(N*0.975,digits=0)
86 D<-rep(0,N)
87 UL<-rep(0,binmax+1)
88 UM<-rep(0,binmax+1)
89 UH<-rep(0,binmax+1)
90 for (j in 1:(binmax+1)){
91     D<-sort(y[1:N, j])
92     Dave<-mean(D)
93     UL[j]<-D[i25]
94     UM[j]<-Dave
95     UH[j]<-D[i75]
96 }
97 xc<-c(0:binmax)
98 SP1<-data.frame(EventRate=xc, frequency=UL, Poisson_D="2.5%")
99 SP2<-data.frame(EventRate=xc, frequency=UM, Poisson_D="mean")
100 SP3<-data.frame(EventRate=xc, frequency=UH, Poisson_D="97.5%")
101 DB2<-rbind(SP3, SP2, SP1)
102 rm(SP1, SP2, SP3, y)
103 #----- ss3
104 lambda<-lambda3
105 binmax<-xmax
106 y<-array(1:N*(binmax+1),dim=c(N,binmax+1))
107 for (i in 1:N){
108     x<-rpois(n, lambda)
109     u<-hist(x, breaks=c(-1:binmax), plot=FALSE)
110     y[i, 1:(binmax+1)]<-u$counts
111 }
112 i25<-round(N*0.025,digits=0)
113 i75<-round(N*0.975,digits=0)
114 D<-rep(0,N)
115 UL<-rep(0,binmax+1)
116 UM<-rep(0,binmax+1)
117 UH<-rep(0,binmax+1)
118 for (j in 1:(binmax+1)){
119     D<-sort(y[1:N, j])

```

```

120     Dave<-mean(D)
121     UL[j]<-D[i25]
122     UM[j]<-Dave
123     UH[j]<-D[i75]
124 }
125 xc<-c(0:binmax)
126 SP1<-data.frame(EventRate=xc, frequency=UL, Poisson_D="2.5%")
127 SP2<-data.frame(EventRate=xc, frequency=UM, Poisson_D="mean")
128 SP3<-data.frame(EventRate=xc, frequency=UH, Poisson_D="97.5%")
129 DB3<-rbind(SP3, SP2, SP1)
130 rm(SP1, SP2, SP3, y)
131 #----- ss4
132 lambda<-lambda4
133 binmax<-xmax
134 y<-array(1:N*(binmax+1), dim=c(N, binmax+1))
135 for (i in 1:N){
136     x<-rpois(n, lambda)
137     u<-hist(x, breaks=c(-1:binmax), plot=FALSE)
138     y[i, 1:(binmax+1)]<-u$counts
139 }
140 i25<-round(N*0.025, digits=0)
141 i75<-round(N*0.975, digits=0)
142 D<-rep(0, N)
143 UL<-rep(0, binmax+1)
144 UM<-rep(0, binmax+1)
145 UH<-rep(0, binmax+1)
146 for (j in 1:(binmax+1)){
147     D<-sort(y[1:N, j])
148     Dave<-mean(D)
149     UL[j]<-D[i25]
150     UM[j]<-Dave
151     UH[j]<-D[i75]
152 }
153 xc<-c(0:binmax)
154 SP1<-data.frame(EventRate=xc, frequency=UL, Poisson_D="2.5%")
155 SP2<-data.frame(EventRate=xc, frequency=UM, Poisson_D="mean")
156 SP3<-data.frame(EventRate=xc, frequency=UH, Poisson_D="97.5%")
157 DB4<-rbind(SP3, SP2, SP1)

```

```

158   rm(SP1, SP2, SP3, y)
159   #---- Make histogram
160   #
161   u1<-hist(ss1,breaks=c(-1:xmax), plot=F)
162   u2<-hist(ss2,breaks=c(-1:xmax), plot=F)
163   u3<-hist(ss3,breaks=c(-1:xmax), plot=F)
164   u4<-hist(ss4,breaks=c(-1:xmax), plot=F)
165   oh1<-u1$counts
166   oh2<-u2$counts
167   oh3<-u3$counts
168   oh4<-u4$counts
169   #
170   OBS<-data.frame(EventRate=xc, oSSI1=oh1, oSSI2=oh2, oSSI3=oh3, oSSI4=oh4)
171   #---- Draw histogram and simulation (Fig.8)
172   pp1<-
ggplot()+layer(data=OBS, aes(x=EventRate, y=oSSI1), geom="bar", stat="identity", position="identity")+laye
r(data=DB1, aes(x=EventRate, y=frequency, group=Poisson_D, color=Poisson_D, size=Poisson_D), geom="line", st
at="identity", position="identity")+annotate("text", label=">=SI1", size=5, x=53, y=15)+labs(x="events/yea
r", y="frequency")+theme(axis.title.x=element_blank())+scale_size_manual(values=c(0.8, 0.8, 0.8))
173   pp2<-
ggplot()+layer(data=OBS, aes(x=EventRate, y=oSSI2), geom="bar", stat="identity", position="identity")+laye
r(data=DB2, aes(x=EventRate, y=frequency, group=Poisson_D, color=Poisson_D, size=Poisson_D), geom="line", st
at="identity", position="identity")+annotate("text", label=">=SI2", size=5, x=20, y=20)+labs(x="events/yea
r", y="frequency")+theme(axis.title.x=element_blank())+scale_size_manual(values=c(0.8, 0.8, 0.8))+xlim(-
0.5, xmax2+1)
174   pp3<-
ggplot()+layer(data=OBS, aes(x=EventRate, y=oSSI3), geom="bar", stat="identity", position="identity")+laye
r(data=DB3, aes(x=EventRate, y=frequency, group=Poisson_D, color=Poisson_D, size=Poisson_D), geom="line", st
at="identity", position="identity")+annotate("text", label=">=SI3", size=5, x=20, y=30)+labs(x="events/yea
r", y="frequency")+theme(axis.title.x=element_blank())+scale_size_manual(values=c(0.8, 0.8, 0.8))+xlim(-
0.5, xmax2+1)
175   pp4<-
ggplot()+layer(data=OBS, aes(x=EventRate, y=oSSI4), geom="bar", stat="identity", position="identity")+laye
r(data=DB4, aes(x=EventRate, y=frequency, group=Poisson_D, color=Poisson_D, size=Poisson_D), geom="line", st
at="identity", position="identity")+annotate("text", label=">=SI4", size=5, x=20, y=60)+labs(x="events/yea
r", y="frequency")+scale_size_manual(values=c(0.8, 0.8, 0.8))+xlim(-0.5, xmax2+1)
176   pp<-cowplot::plot_grid(pp1, pp2, pp3, pp4, labels=c("a", "b", "c", "d"), align="v", nrow=4)
177   print(pp)

```

読み込みデータファイル” kochiHonmachi.xlsx” は次の並びになっている。1 行目がヘッダーで、2 行目の左から年、震度 1 の回数、震度 2 の回数、...、震度 7 の回数となっている。

	A	B	C	D	E	F	G	H	I	J
1	期間	震度 1	震度 2	震度 3	震度 4	震度 5 弱	震度 5 強	震度 6 弱	震度 6 強	震度 7
2	1923年	1	1	2	0	0	0	0	0	0
3	1924年	0	3	0	0	0	0	0	0	0

(2) 図 14 の作図スクリプト

```

1   rm(list=ls())
2   library(tidyverse)
3   library(cowplot)
4   library(readxl)
5   #---- read data
6   # EQ$OriginTime, Epicenter, Lat, Lon, Depth, Magnitude, MaxSI, SI
7   EQ<-read_excel("HonmachiEQ3.xlsx", sheet=1, col_names=FALSE, skip=2)
8   colnames(EQ)<-c("OriginTime", "Epicenter", "Lat", "Lon", "Depth", "Magnitude", "MaxSI", "SI")
9   #---- Draw time series (Fig.14a)
10  num<-nrow(EQ)
11  addDB1<-data.frame(Magnitude=as.numeric(str_sub(EQ$Magnitude, start=2, end=4)))
12  addDB2<-data.frame(ymin=rep(0, num))
13  subEQ<-EQ[, c("OriginTime")]
14  EQDB<-cbind(subEQ, addDB1, addDB2)
15  gg<-ggplot(data=EQDB)+geom_point(aes(x=OriginTime, y=Magnitude))
16  gg<-gg+geom_linerange(mapping=aes(x=OriginTime, ymin=0, ymax=Magnitude))
17  gg<-gg+labs(x="year")
18  #---- calc. occurrence interval
19  dt<-diff(EQ$OriginTime)
20  units(dt)<-"hours"
21  DT<-as.numeric(dt)/(24*365)
22  #---- Set parameters
23  N <- 10000
24  wbin <- 1
25  set.seed(190403)
26  #---- Sim. Exponential Dist.
27  binmax<-round(10*max(DT)/wbin)*wbin
28  n<-num-1

```

```

29     lambda<-mean(DT)
30     rate<-1/lambda
31     #
32     y<-array(1:N*(binmax), dim=c(N, binmax))
33     for (i in 1:N){
34         x<-rexp(n, rate)
35         u<-hist(x, breaks=seq(0, binmax, wbin), right=FALSE, plot=FALSE)
36         y[i, 1:(binmax/wbin)]<-u$counts
37     }
38     i25<-round(N*0.025, digits=0)
39     i75<-round(N*0.975, digits=0)
40     D<-rep(0, N)
41     UL<-rep(0, binmax/wbin)
42     UM<-rep(0, binmax/wbin)
43     UH<-rep(0, binmax/wbin)
44     for (j in 1:(binmax/wbin)){
45         D<-sort(y[1:N, j])
46         UL[j]<-D[i25]
47         UM[j]<-mean(D)
48         UH[j]<-D[i75]
49     }
50     rm(y)
51     #---- Draw histogram of obs data (Fig14b)
52     obsDT <- data.frame(DT)
53     #
54     ghobs <- ggplot(obsDT, aes(x=DT))
55     ghobs <- ghobs+geom_histogram(breaks=seq(0, 45, wbin), closed='left')+ylim(0, 50)
56     #---- Draw Upper, mean and Lower points
57     xc<-seq(wbin/2, binmax, by=wbin)
58     DB1<-data.frame(Interval=xc, frequency=UL, Exponential_D="2.5%")
59     DB2<-data.frame(Interval=xc, frequency=UM, Exponential_D="mean")
60     DB3<-data.frame(Interval=xc, frequency=UH, Exponential_D="97.5%")
61     DB<-rbind(DB3, DB2, DB1)
62     #
63     h1<-ghobs+layer(geom="line", stat="identity", subset(DB, Exponential_D %in%
c("97.5%", "mean", "2.5%")), aes(x=Interval, y=frequency, group=Exponential_D, colour=Exponential_D, size=Ex
ponential_D), position="identity")+scale_size_manual(values=c(0.8, 0.8, 0.8))
64     h1<-h1+xlim(0, 45)+ylim(0, 50)

```



```

65     h1<-h1+labs(x="Occurrence interval, year")+labs(y="frequency")+annotate("text", label=">=SI3 at
Honmachi (1923-2017)", x=15, y=25, size=5)+theme(legend.position=c(0.8, 0.6), legend.justification=c(0,1))
66     pp<-cowplot::plot_grid(gg, h1, labels=c("a", "b"), align="v", nrow=2)
67     print(pp)
68     #----- QQ plot
69     #EQDT<-data.frame(DT=DT)
70     #gg<-
ggplot(data=EQDT, mapping=aes(sample=DT))+stat_qq_line(distribution=qexp)+stat_qq(distribution=qexp)+l
abs(x="Exp. theoretical quantiles", y="Occurrence interval quantiles")
71     #print(gg)

```

読み込みデータファイル” HonmachiEQ3.xlsx” は次の並びになっている。1 行目及び 2 行目がヘッダ行で、3 行目の左から地震の発生日時、震央地名、緯度、経度、深さ、M(マグニチュード)、最大震度、検索対象最大震度となっている。図 14a の作図には、地震の発生日時と M(マグニチュード)の項目だけを使用する。

	A	B	C	D	E	F	G	H
1	地震の発生日時	震央地名	緯度	経度	深さ	M	最大震度	検索対象
2								最大震度
3	1923/12/05 08:40:25.9	土佐湾	33° 14.0' N	133° 53.7' E	16km	M6.3	4	3
4	1923/12/12 13:04:06.3	瀬戸内海中 部	34° 08.1' N	133° 21.9' E	27km	M5.4	3	3
5	1927/03/07 18:27:39.2	京都府北部	35° 37.9' N	134° 55.8' E	18km	M7.3	6	3