Assignment I - CompStat2023

Aldo Giovanni e Giacomo

Anuja Saira Abraham 5204982, Alessia Marzotti 5108443, Miriam Mercuri 5207057

Set up

Simulation problem: `RISIKO!`

Point a

Compute $P[\max(X_1, X_2) > Y_1]$.

We know that $Z = max(X_1, X_2) \in (1, 2, 3, 4, 5, 6)$ and that $P(Z = z) = (2(z-1)+1)/36$. In other word we want to find $P(Z > y) = 1 - P(Z \leq y)$. $P(Z \leq y) =$
$P(Z = 1 \cap y = 1)+$
$P(Z = 1 \cap y = 2) + P(Z = 2 \cap y = 2)+$
$P(Z = 1 \cap y = 3) + P(Z = 2 \cap y = 3) + P(Z = 3 \cap y = 3)+$
$P(Z = 1 \cap y = 4) + P(Z = 2 \cap y = 4) + P(Z = 3 \cap y = 4) + P(Z = 4 \cap y = 4)+$
$P(Z = 1 \cap y = 5) + P(Z = 2 \cap y = 5) + P(Z = 3 \cap y = 5) + P(Z = 4 \cap y = 5) + P(Z = 5 \cap y = 5)+$
$P(Z = 1 \cap y = 6) + P(Z = 2 \cap y = 6) + P(Z = 3 \cap y = 6) + P(Z = 4 \cap y = 6) + P(Z = 5 \cap y = 6) + P(Z = 6 \cap y = 6)$

So we have $P(Z \leq y) = 6\frac{1}{36}\frac{1}{6} + 5\frac{3}{36}\frac{1}{6} + 4\frac{5}{36}\frac{1}{6} + 3\frac{7}{36}\frac{1}{6} + 2\frac{9}{36}\frac{1}{6} + \frac{11}{36}\frac{1}{6} \approx 0,421$ and

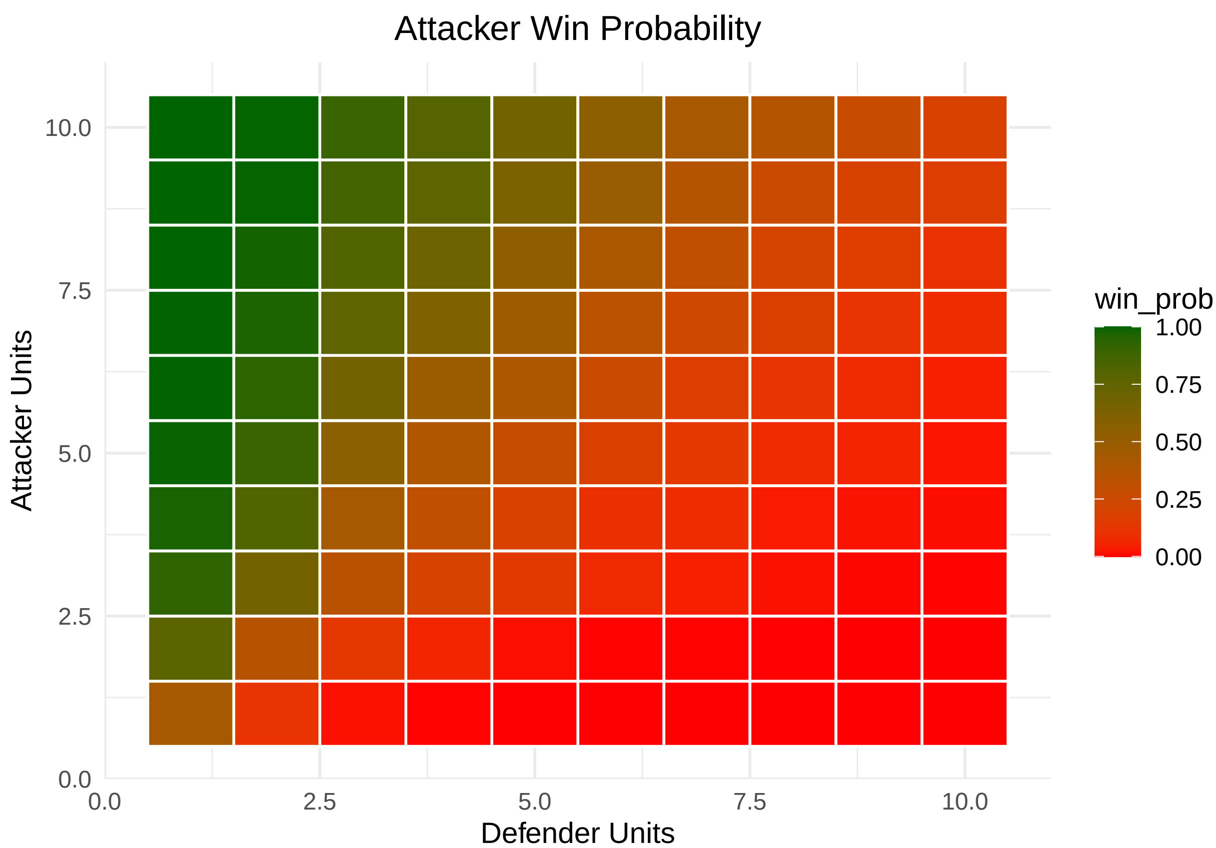
$$P(Z > y) = 1 - 0,421 = 0,579$$

Point b-c-d

Here there are some code to simulate a generic `Risiko!` game for different values of competing units.

```
library(ggplot2)

set.seed(123)
combat_round <- function(att_units,def_units,sim=10000) {
  Results = rep(NA,sim)
  AS<-att_units
  DS<-def_units
  for(i in 1:sim){
    while(def_units>0 & att_units>0){
      Dnum <- sort(sample(1:6, min(def_units,3),replace = TRUE),decreasing = T)
      Anum <- sort(sample(1:6, min(att_units,3),replace = TRUE),decreasing = T)
      for (j in 1:min(length(Dnum),length(Anum))){
        if(Anum[j]>Dnum[j]){
          def_units<-def_units-1
        }
        else{
          att_units<-att_units-1
        }
      }
```

```r
    }
    Results[i]<- ifelse(att_units>0,1,0)
    att_units<-AS
    def_units<-DS
  }
  return(mean(Results))
}

prob_att_win <- function(){
  res<-sapply(1:10, function(x) {
  sapply(1:10, function(y) {
    combat_round(y,x,1000)
    })
  })
  return(res)
}
# Print the result
#print(result)


att_prob<- data.frame(
  attacker_unit<- rep(1:10,10),
  defender_unit<- rep(1:10,each= 10),
  win_prob<- as.vector(prob_att_win())
)


#go through outer
ggplot(att_prob,aes(x=defender_unit,y=attacker_unit,fill=win_prob))+
  geom_tile()+
  scale_fill_gradient(low="red",high="darkgreen")+
  labs(title = "Attacker Win Probability",x="Defender Units",y="Attacker Units")+
  geom_tile(color = "white",lwd = 0.5,linetype = 1)+
  theme_minimal()+
  theme(plot.title = element_text(hjust = 0.5))
```




```
print(combat_round(def_units=1,att_units=2,sim=10000))
```

```
## [1] 0.7561
```

Monte Carlo simulations I

Point a

$E(X) = \int_a^b \frac{x}{b-a}dx = \frac{b^2-a^2}{2(b-a)}$

$E(Z) = E(\sum_{i=1}^{12} U_i) = \sum_{i=1}^{12} E(U_i)$ independent $= 12E(U_i)$ identically distributed $12\frac{\frac{1}{2}^2-(-\frac{1}{2})^2}{2(\frac{1}{2}-(-\frac{1}{2}))} = 0$

$V(Z) = E[(X-E(X))^2] = \int_a^b (x-\frac{a+b}{2})^2\frac{dx}{b-a} = \frac{(b-a)^2}{12}$

$V(Z) = V(\sum_{i=1}^{12} U_i) = \sum_{i=1}^{12} V(U_i)$ variable independent $= 12V(U_i)$ U_i identically distributed $= 12\frac{(\frac{1}{2}-(-\frac{1}{2}))^2}{12}$ $= 12\frac{(\frac{1}{2}+\frac{1}{2})^2}{12} = 1$

```
set.seed(13)



normal_functn_gen = function(sim){
  n<-12
  Uz<-runif(sim*n,min = -0.5,max = 0.5)
  Uz<-matrix(Uz,nrow=N,ncol=n)
  Z<-apply(Uz,1,sum)
```

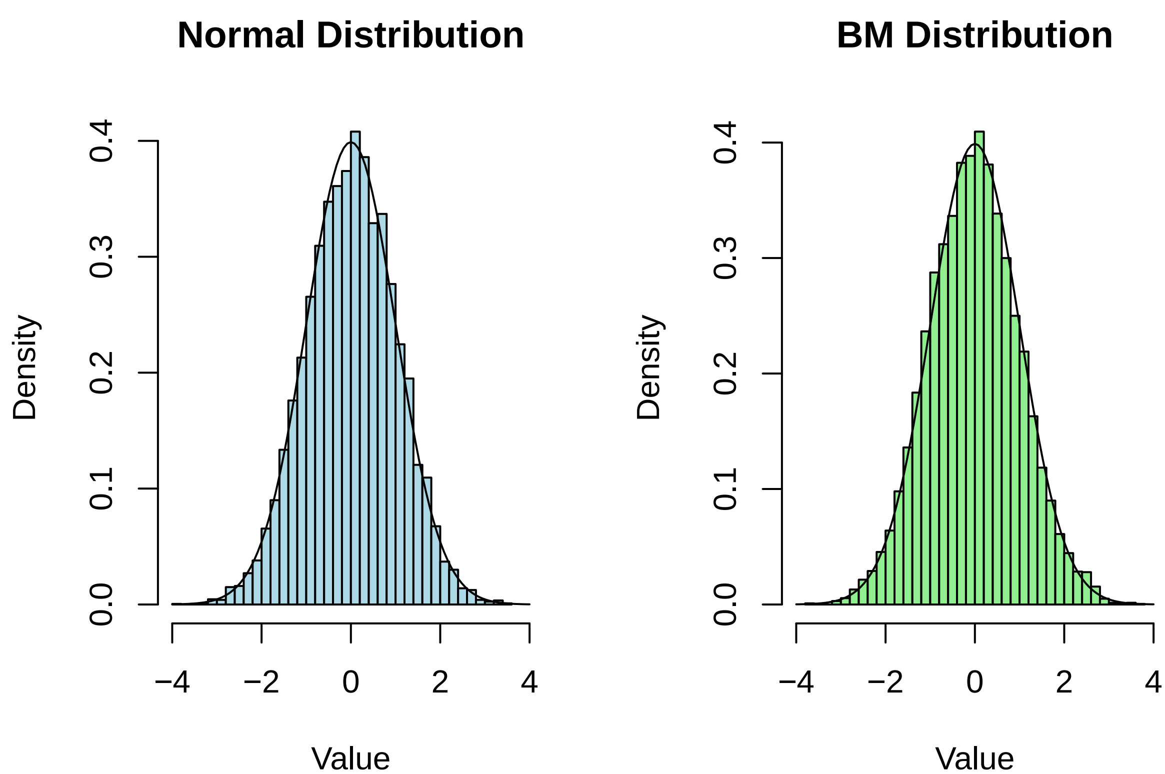
```
  return(Z)
}

N <- 10000

gen_norm<-normal_functn_gen(sim=10000)
mean(gen_norm)
```

```
## [1] 0.01682814
```

```
var(gen_norm)
```

```
## [1] 0.9994385
```

```
U1 <- runif(N)
U2 <- runif(N)
X1 <- sqrt( -2*log(U1) )*cos(2*pi*U2)

par(mfrow=c(1,2))
hist(gen_norm, main="Normal Distribution", xlab="Value",col="lightblue" ,breaks=30,freq = F,xlim=c(-4,4
curve(dnorm(x,0,1),add=T)
hist(X1, main="BM Distribution", xlab="Value",col="lightgreen" ,breaks=30,freq = F,xlim=c(-4,4))
curve(dnorm(x,0,1),add=T)
```



Monte Carlo simulations II

Point a

The Pareto distribution is defined by a density $f(x;\gamma) = \gamma x^{-(\gamma+1)}$ over $(1;+\infty)$, with $\gamma > 0$.

It can be generated as the $-\frac{1}{\gamma}$ power of a uniform r.v.

Cumulative distribution function of Pareto distribution $\int_1^x \gamma z^{-(\gamma+1)}dz = \gamma \int_1^x z^{-1-\gamma}dz = -[x^{-\gamma} - 1^{-\gamma}] = -x^{-\gamma} + 1 = 1 - (\frac{1}{x})^{\gamma}$

We will use the following theorem: if $X \sim F(x)$ then $U = F(x) \sim U(0,1)$ $F(X) = 1 - (\frac{1}{x})^{\gamma} = U$ $(1-U)^{-\frac{1}{\gamma}} = (x^{\gamma})^{-\frac{1}{\gamma}}$ $x = (1-U)^{-\frac{1}{\gamma}} = U^{-\frac{1}{\gamma}}$

Pont b

Reference: Wikipedia The Pareto distribution is related to the exponential distribution as follows. If X is Pareto-distributed with minimum x_m and index α, then $Y = log(\frac{X}{x_m})$ is exponentially distributed with rate parameter α.

$Y = log(\frac{X}{x_m}) \sim Exp(\gamma)$ $Y = log(X) \sim Exp(\gamma)$

##Point c

Implement two samplers, one for X and one for Y . Plot the histogram and the density and comment on the results exploring different values of Y.

MC integration

Point a

We write the probability of a standard Normal r.v. X as an integral thanks to its probability density function.

$P(X > 20) = \int_{20}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

Reference: Wikipedia

The crude Monte Carlo estimation of this quantity is deemed to fail because the region of integration is so far out in the tails of the standard normal distribution.

Point b

Rewrite the integral employing the change of variable $Y = \frac{1}{X}$.

$dy = -\frac{1}{x^2}dx$
$dx = -x^2 dy = -\frac{1}{y^2}dy$
$y = Y_{20} = 0,05$
$y = Y_{\infty} = 0$

So $\int_{20}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$
$= \int_{0,05}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2y^2}} - \frac{1}{y^2} dy$
$= \int_{0}^{0,05} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2y^2}} \frac{1}{y^2} dy$