**INTRODUCTION**

**Consider a random experiment and an event A. Suppose that it is possible to repeat the experiment a large number of times under essentially similar conditions. Denote f(A) as the number of occurrences of A in n repetitions.**

**Then f(A)/n is called the relative frequency of A. A kind of regularity is observed when a large number of repetitions is considered together and it is an observed fact that the relative frequencies stabilize to a certain value as n becomes large. This kind of regularity in a random experiment is known as *STATISTICAL REGULARITY*. The limiting value of f(A)/n as**

**n ∞, is called the probability of the event A, provided the limit exists. Now, it is not possible to repeat the experiment manually under essentially similar conditions. For that reason, we perform the experiment using *SIMULATION* technique with the help of statistical programming language ‘ R ’**

**WHAT IS SIMULATION?**

**Simulation is a way to model random events, such that simulated outcomes closely match real-world outcomes.**

**By observing simulated outcomes, researchers gain insight on the real world.**

**WHY DO WE USE SIMULATION?**

**Some situations do not lend themselves to precise mathematical treatment. Others may be difficult, time-consuming, or expensive to analyze. In these situations, simulation may approximate real-world results; yet, require less time, effort, and/or money than other approaches.**

EXAMPLE 1

**A population consists of 200 people, among them there are 125 males and 75 females. A sample of size 30 is drawn without replacement. Find the probability that the sample will contain at least 10 females.**

**BY CLASSICAL DEFINITION OF PROBABILTY:**

**Given, that the size of the population = 200**

**No. of males = 125**

**No. of females = 75**

**Assuming that the events of this random experiment are mutually exclusive, exhaustive and equally likely.**

**According to the problem, the sample is drawn without replacement. Hence, all possible outcomes of this experiment will be 200C30 .**

**Now let A be the event denoting that the sample will contain at least 10 females.**

**So, clearly favourable cases for the event A will be , 75Ci × 125C30-i .**

**By classical definition of probability,**

**P(A)= (75Ci × 125C30-i ) / 200C30**

**= 0.76 (approx.)**

**BY SIMULATION TECHNIQUE:**

**rm(list=ls())**

**#to check whether the sample contains at least 10 females**

**countSample=function(x)**

**{**

**count=0**

**i=1**

**while(i<=30){**

**if(x[i]==1){**

**count=count+1**

**}**

**i=i+1**

**}**

**if(count>=10){**

**return (TRUE)**

**}else{**

**return (FALSE)**

**}**

**}**

**sampleGenerate=function(n)**

**{**

**i=1**

**count=0**

**#we denote the males by 0 and females by 1**

**x=c(rep(0,125),rep(1,75))**

**while(i<=n){**

**#drawing a random sample of size 30 from a population of size 200 y=sample(x,size=30,replace=FALSE)**

**if(countSample(y)){**

**count=count+1**

**}**

**i=i+1**

**}**

**p=count/n**

**return (p)**

**}**

**draw=function()**

**{**

**w=1**

**x=rep(0,100)**

**y=rep(0,100)**

**i=1000**

**while(i<=100000){**

**x[w]=sampleGenerate(i)**

**y[w]=i**

**w=w+1**

**i=i+1000**

**}**

**plot(y,x,type="l",xlab="no. of repetitions**

**of the experiment",ylab="relative frequency")**

**}**



|  |  |
| --- | --- |
| THEORETICAL VALUE | SIMULATED VALUE |
| 0.76 | **0.758** |

EXAMPLE 2

**An unbiased die is thrown 3 times. Let a, b and c denotes the outcome of the first, second and third throws respectively. Consider the equation ax2+bx+c=0. Find the probability that the roots of the equation will be real.**

**BY CLASSICAL DEFINITION OF PROBABILTY:**

**In the problem an unbiased die is thrown three times and the outcomes are observed.**

**Assuming that the events of this random experiment are mutually exclusive, exhaustive and equally likely.**

**Now, for real roots of the equation,**

**ax2+bx+c = 0**

**b2 – 4ac ≥ 0 (\*)**

**Defining (\*) as an event B, where a, b and c are the first, second and third outcome of the die respectively.**

**The all possible cases along with its favourable cases are given in the table below. (The favourable cases are shown in red )**

**Clearly, from the table there are 216 all possible cases and 43 outcomes are favourable to the event B.**

**Hence, by classical definition of probability,**

**P(B) = 43/216 = 0.2 (approx)**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| a | b | c | b2-4ac | a | b | c | b2-4ac | a | b | c | b2-4ac | a | b | c | b2-4ac | a | b | c | b2-4ac | a | b | c | b2-4ac |
| 1 | 1 | 1 | -3 | 2 | 1 | 1 | -7 | 3 | 1 | 1 | -11 | 4 | 1 | 1 | -15 | 5 | 1 | 1 | -19 | 6 | 1 | 1 | -23 |
| 1 | 1 | 2 | -7 | 2 | 1 | 2 | -15 | 3 | 1 | 2 | -23 | 4 | 1 | 2 | -31 | 5 | 1 | 2 | -39 | 6 | 1 | 2 | -47 |
| 1 | 1 | 3 | -11 | 2 | 1 | 3 | -23 | 3 | 1 | 3 | -35 | 4 | 1 | 3 | -47 | 5 | 1 | 3 | -59 | 6 | 1 | 3 | -71 |
| 1 | 1 | 4 | -15 | 2 | 1 | 4 | -31 | 3 | 1 | 4 | -47 | 4 | 1 | 4 | -63 | 5 | 1 | 4 | -79 | 6 | 1 | 4 | -95 |
| 1 | 1 | 5 | -19 | 2 | 1 | 5 | -39 | 3 | 1 | 5 | -59 | 4 | 1 | 5 | -79 | 5 | 1 | 5 | -99 | 6 | 1 | 5 | -119 |
| 1 | 1 | 6 | -23 | 2 | 1 | 6 | -47 | 3 | 1 | 6 | -71 | 4 | 1 | 6 | -95 | 5 | 1 | 6 | -119 | 6 | 1 | 6 | -143 |
| 1 | **2** | **1** | **0** | 2 | 2 | 1 | -4 | 3 | 2 | 1 | -8 | 4 | 2 | 1 | -12 | 5 | 2 | 1 | -16 | 6 | 2 | 1 | -20 |
| 1 | 2 | 2 | -4 | 2 | 2 | 2 | -12 | 3 | 2 | 2 | -20 | 4 | 2 | 2 | -28 | 5 | 2 | 2 | -36 | 6 | 2 | 2 | -44 |
| 1 | 2 | 3 | -8 | 2 | 2 | 3 | -20 | 3 | 2 | 3 | -32 | 4 | 2 | 3 | -44 | 5 | 2 | 3 | -56 | 6 | 2 | 3 | -68 |
| 1 | 2 | 4 | -12 | 2 | 2 | 4 | -28 | 3 | 2 | 4 | -44 | 4 | 2 | 4 | -60 | 5 | 2 | 4 | -76 | 6 | 2 | 4 | -92 |
| 1 | 2 | 5 | -16 | 2 | 2 | 5 | -36 | 3 | 2 | 5 | -56 | 4 | 2 | 5 | -76 | 5 | 2 | 5 | -96 | 6 | 2 | 5 | -116 |
| 1 | 2 | 6 | -20 | 2 | 2 | 6 | -42 | 3 | 2 | 6 | -68 | 4 | 2 | 6 | -92 | 5 | 2 | 6 | -116 | 6 | 2 | 6 | -140 |
| 1 | **3** | **1** | **5** | **2** | **3** | **1** | **1** | 3 | 3 | 1 | -3 | 4 | 3 | 1 | -7 | 5 | 3 | 1 | -11 | 6 | 3 | 1 | -15 |
| 1 | **3** | **2** | **1** | 2 | 3 | 2 | -7 | 3 | 3 | 2 | -15 | 4 | 3 | 2 | -23 | 5 | 3 | 2 | -31 | 6 | 3 | 2 | -39 |
| 1 | 3 | 3 | -3 | 2 | 3 | 3 | -15 | 3 | 3 | 3 | -27 | 4 | 3 | 3 | -39 | 5 | 3 | 3 | -51 | 6 | 3 | 3 | -63 |
| 1 | 3 | 4 | -7 | 2 | 3 | 4 | -23 | 3 | 3 | 4 | -39 | 4 | 3 | 4 | -55 | 5 | 3 | 4 | -71 | 6 | 3 | 4 | -87 |
| 1 | 3 | 5 | -11 | 2 | 3 | 5 | -31 | 3 | 3 | 5 | -51 | 4 | 3 | 5 | -71 | 5 | 3 | 5 | -91 | 6 | 3 | 5 | -111 |
| 1 | 3 | 6 | -15 | 2 | 3 | 6 | -39 | 3 | 3 | 6 | -63 | 4 | 3 | 6 | -87 | 5 | 3 | 6 | -111 | 6 | 3 | 6 | -135 |
| 1 | **4** | **1** | **12** | **2** | **4** | **1** | **8** | **3** | **4** | **1** | **4** | **4** | **4** | **1** | **0** | 5 | 4 | 1 | -4 | 6 | 4 | 1 | -8 |
| 1 | **4** | **2** | **8** | **2** | **4** | **2** | **0** | 3 | 4 | 2 | -8 | 4 | 4 | 2 | -16 | 5 | 4 | 2 | -24 | 6 | 4 | 2 | -32 |
| 1 | **4** | **3** | **4** | 2 | 4 | 3 | -8 | 3 | 4 | 3 | -20 | 4 | 4 | 3 | -32 | 5 | 4 | 3 | -44 | 6 | 4 | 3 | -56 |
| 1 | **4** | **4** | **0** | 2 | 4 | 4 | -16 | 3 | 4 | 4 | -32 | 4 | 4 | 4 | -48 | 5 | 4 | 4 | -64 | 6 | 4 | 4 | -80 |
| 1 | 4 | 5 | -4 | 2 | 4 | 5 | -24 | 3 | 4 | 5 | -44 | 4 | 4 | 5 | -64 | 5 | 4 | 5 | -84 | 6 | 4 | 5 | -104 |
| 1 | 4 | 6 | -8 | 2 | 4 | 6 | -32 | 3 | 4 | 6 | -56 | 4 | 4 | 6 | -80 | 5 | 4 | 6 | -104 | 6 | 4 | 6 | -128 |
| 1 | **5** | **1** | **21** | **2** | **5** | **1** | **17** | **3** | **5** | **1** | **13** | **4** | **5** | **1** | **9** | **5** | **5** | **1** | **5** | **6** | **5** | **1** | **1** |
| 1 | **5** | **2** | **17** | **2** | **5** | **2** | **9** | **3** | **5** | **2** | **1** | 4 | 5 | 2 | -7 | 5 | 5 | 2 | -15 | 6 | 5 | 2 | -23 |
| 1 | **5** | **3** | **13** | **2** | **5** | **3** | **1** | 3 | 5 | 3 | -11 | 4 | 5 | 3 | -23 | 5 | 5 | 3 | -35 | 6 | 5 | 3 | -47 |
| 1 | **5** | **4** | **9** | 2 | 5 | 4 | -7 | 3 | 5 | 4 | -23 | 4 | 5 | 4 | -39 | 5 | 5 | 4 | -55 | 6 | 5 | 4 | -71 |
| 1 | **5** | **5** | **5** | 2 | 5 | 5 | -15 | 3 | 5 | 5 | -35 | 4 | 5 | 5 | -55 | 5 | 5 | 5 | -75 | 6 | 5 | 5 | -95 |
| 1 | **5** | **6** | **1** | 2 | 5 | 6 | -23 | 3 | 5 | 6 | -47 | 4 | 5 | 6 | -71 | 5 | 5 | 6 | -95 | 6 | 5 | 6 | -119 |
| 1 | **6** | **1** | **32** | **2** | **6** | **1** | **28** | **3** | **6** | **1** | **24** | **4** | **6** | **1** | **20** | **5** | **6** | **1** | **16** | **6** | **6** | **1** | **12** |
| 1 | **6** | **2** | **28** | **2** | **6** | **2** | **20** | **3** | **6** | **2** | **12** | **4** | **6** | **2** | **4** | 5 | 6 | 2 | -4 | 6 | 6 | 2 | -12 |
| 1 | **6** | **3** | **24** | **2** | **6** | **3** | **12** | **3** | **6** | **3** | **0** | 4 | 6 | 3 | -12 | 5 | 6 | 3 | -24 | 6 | 6 | 3 | -36 |
| 1 | **6** | **4** | **20** | **2** | **6** | **4** | **4** | 3 | 6 | 4 | -12 | 4 | 6 | 4 | -28 | 5 | 6 | 4 | -44 | 6 | 6 | 4 | -60 |
| 1 | **6** | **5** | **16** | 2 | 6 | 5 | -4 | 3 | 6 | 5 | -24 | 4 | 6 | 5 | -44 | 5 | 6 | 5 | -64 | 6 | 6 | 5 | -84 |
| 1 | **6** | **6** | **12** | 2 | 6 | 6 | -12 | 3 | 6 | 6 | -36 | 4 | 6 | 6 | -60 | 5 | 6 | 6 | -84 | 6 | 6 | 6 | -108 |

**BY SIMULATION TECHNIQUE:**

**rm(list=ls())**

**#to find whether the roots of the equation a\*x^2+b\*x+c=0 are real**

**isReal=function(a,b,c)**

**{**

**#calculating the discriminant**

**d= b^2-4\*a\*c**

**if(d>=0)**

**{**

**return (TRUE)**

**}**

**else{**

**return (FALSE)**

**}**

**}**

**#to calculate the relative frequency**

**probability=function(n)**

**{**

**i=1**

**count=0**

**while(i<=n)**

**{**

**#throwing a dice three times**

**a=sample(1:6,size=1)**

**b=sample(1:6,size=1)**

**c=sample(1:6,size=1)**

**if(isReal(a,b,c))**

**{**

**count=count+1**

**}**

**i=i+1**

**}**

**#calculating the relative frequency**

**p=count/n**

**return (p)**

**}**

**draw=function()**

**{**

**x=rep(0,100)**

**y=rep(0,100)**

**w=1**

**i=1000**

**while(i<=100000)**

**{**

**x[w]=probability(i)**

**y[w]=i**

**i=i+1000**

**w=w+1**

**}**

**plot(y,x,type="l",xlab="no. of repetitions of the**

** experiment",ylab="relative frequency")**

**}**

|  |  |
| --- | --- |
| THEORETICAL VALUE | SIMULATED VALUE |
| 0.20 | **0.201** |

EXAMPLE 3

**Imagine that a chess player, whom we’ll call A, has been challenged to a curious sort of match by two of his competitors, whom we’ll call B and C. A is challenged to play three sequential games alternating between B and C, with the first game being with the player A’s choice. That is, A could play either BCB (we’ll call this sequence 1) or CBC (and this will be sequence 2). From experience with B and C, A knows that C is the stronger player (the tougher for A to defeat). Indeed, from experience, A attaches probabilities *p* and *q* to his likelihood of winning any particular game against B and C, respectively, where *q* < *p*. The rule of this peculiar match is that to win the challenge (i.e., to win the match) A must win two games in a row. (This means in particular, that even if A wins the first and third games - two out of three games – A still loses the match) So, which sequence should A choose to give himself the best chance of winning the match?**

**BY CLASSICAL DEFINITION OF PROBABILTY:**

**For A to win the match, there are just two ways to do so. Either he wins the first two games (and the third game is then irrelevant) or he loses the first game and wins the final two games. Let P1 and P2 be the probability A wins the match playing sequence 1 and sequence 2 respectively. Then, making the usual assumption of independence from game to game, we have,**

**For sequence 1 : P1 = *pq* + (1-*p*)*qp***

**For sequence 2 : P2 = *qp* + (1-*q*)*pq***

**Therefore,**

**P2 – P1 = [*qp* + (1-*q*)*pq*] – [*pq* + (1-*p*)*qp*]**

**= (1-*q*)*pq* – (1-*p*)*qp***

**= *pq* (*p* - *q*) > 0 ,**

**because *q* < *p* . So, sequence 2 always, for any *p* > *q* ,gives the greater probability for A winning the challenge match, even though sequence 2 requires A to play his stronger opponent twice.**

**BY SIMULATION TECHNIQUE:**

**rm(list=ls())**

**playGame=function()**

**{**

**p=as.numeric(readline("enter p : "))**

**q=as.numeric(readline("enter q : "))**

**prob1=c(p,q,p)**

**prob2=c(q,p,q)**

**i=k=l=1**

**count1=count2=0**

**win1=c(0,0,0)**

**win2=c(0,0,0)**

**result=c(0,0,0)**

**while(i<=100000)**

**{**

**result=runif(3,0,1)**

**while(k<=3){**

**if(result[k]<prob1[k])**

**{**

**win1[k]=1**

**}**

**k=k+1**

**}**

**if(((win1[1]==1)&&(win1[2]==1))||**

**((win1[2]==1)&&(win1[3]==1)))**

**{**

**count1=count1+1**

**}**

**k=1**

**while(l<=3){**

**if(result[l]<prob2[l])**

**{**

**win2[l]=1**

**}**

**l=l+1**

**}**

**if(((win2[1]==1)&&(win2[2]==1))||**

**((win2[2]==1)&&(win2[3]==1)))**

**{**

**count2=count2+1**

**}**

**l=1**

**win1=c(0,0,0)**

**win2=c(0,0,0)**

**i=i+1**

**}**

**x=c((count1/100000),(count2/100000))**

**return (x)**

**}**

***The following table compares the estimates of P1 and P2 produced by simulation technique, for some selected values of p and q, to the numbers produced by the theoretical calculations we calculated earlier.***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | **THEORETICAL** | | **SIMULATED** | |
| **p** | **q** | **P1** | **P2** | **P1** | **P2** |
| 0.9 | **0.8** | 0.792 | 0.864 | 0.794 | 0.866 |
| **0.9** | **0.4** | 0.396 | 0.576 | 0.397 | 0.578 |
| **0.6** | **0.2** | 0.192 | 0.204 | 0.191 | 0.203 |
| **0.4** | **0.3** | 0.064 | 0.076 | 0.063 | 0.075 |
| **0.4** | **0.1** | 0.168 | 0.216 | 0.166 | 0.216 |

**CONCLUSION**

**Thus we see that with the help of *SIMULATION* technique, we can perform the random experiment a large number of times under essentially similar conditions, which is much less time consuming and lastly we can compute the probability of a particular event with much ease.**