

# STAT 527 HW 3

Satoshi Ido (ID: 34788706)

5 February 2023

1

(10 points) Money growth. Write an R function to calculate the amount of money,  $nt$ , in a fixed interest investment at any time  $t$  in the future, starting with  $n0$  dollars at  $100 \cdot i\%$  per year interest.

The equation is  $nt = n0(1+i)^t$ .

Besides  $n0$  and  $i$ , the function should take as an argument a whole vector of times and return a whole vector of corresponding dollar amounts. Evaluate your function with appropriate experimental inputs.

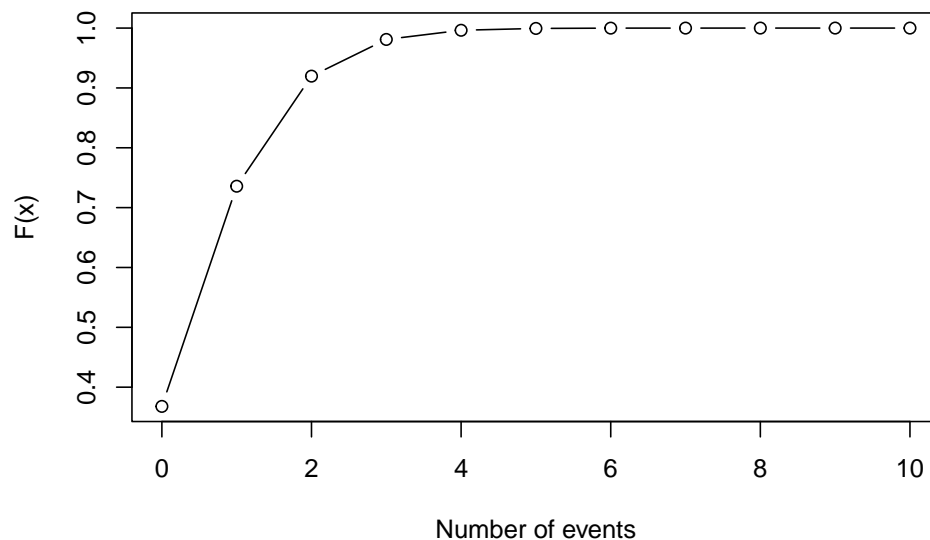
```
nt <- function(n0, i, t) {  
  out <- n0 * (1 + i)^t  
  return(out)  
}  
  
# This function seems work right  
nt(100, 0.05, 10)
```

```
## [1] 162.8895
```

2 (10 points) Poisson distribution. A discrete random variable  $X$  is said to have a Poisson distribution, with parameter  $\lambda > 0$ , if it has a probability mass function given by:  $f(k; \lambda) = \Pr(X = k) = (\lambda^k * e^{-\lambda}) / k!$  where  $k \geq 0$ . The cumulative distribution function (CDF) of Poisson distribution evaluated at  $X = x$  is calculated as  $\Pr(X \leq x)$ . Write an R function that takes  $\lambda$  and  $x$  as inputs and returns the CDF value. Pick  $\lambda=1$  and evaluate your function at  $x = 0, 1, \dots, 10$ . Plot your results and set an appropriate title for your plot. Hint: you may find the R functions `exp()` and `factorial()` being useful.

```
cdf <- function(lambda, x) {  
  a <- 0:x  
  pmf <- lambda ^ a * exp(lambda * (-1)) / factorial(a)  
  out <- cumsum(pmf)  
  return(out)  
}  
a <- 0:10  
plot(a, cdf(1, 10), xlab = "Number of events", ylab = "F(x)",  
     main = "The plot of CDF for Poisson distribution", type = "b")
```

**The plot of CDF for Poisson distribution**



3

(10 points) Hardy-Weinberg frequencies. Remember from basic genetics that in humans and many animals a gene inherited from the mother and a gene inherited from the father combine to determine a trait inherited by an offspring. If there are two variants of a gene, labeled, say, A and B, in a population, then the possible genotypes of individuals in the population are AA, AB, or BB. Suppose we model the formation of a new generation as follows. Males and females put all their gametes or sex cells, each of which carries either an A gene or a B gene, into a big urn, and new individuals are made by drawing out two gametes at random. Suppose the proportion (or fraction) of gametes with A genes in the urn was  $p$ , and so the proportion of gametes with B genes was  $1 - p$ . These two fractions are called the gene frequencies of A and B in the population. Draw out two gametes at random. The probability of drawing two As and forming an AA genotype individual is  $p^2$ , the probability of an AB individual is  $2p(1 - p)$  (an AB individual can be formed by drawing AB or BA), and the probability of a BB individual is  $(1 - p)^2$ . These genotype frequencies, produced theoretically by “random mating,” are called the Hardy-Weinberg frequencies. Write an R function to take any gene frequency for gene A (where the frequency is between 0 and 1) and calculate the Hardy-Weinberg genotype frequencies. The function should take as an argument a single number and return a vector of three quantities. Evaluate your function with appropriate experimental inputs.

```
gene <- function(p) {
  AA <- p^2
  AB <- 2 * p * (1 - p)
  BB <- (1 - p)^2
  return(c(AA, AB, BB))
}

# I suppose the proportion (or fraction) of gametes with A genes in the urn (p) is 0.2
gene(0.2)
```

```
## [1] 0.04 0.32 0.64
```