

Introduction to R — Fall 2007

Homework Set # 1

Begin the homework session by executing the following command in *R*:

```
source("http://www.nd.edu/~steve/Rcourse/hmwrkData/hmwrk1Data.R")
```

This will define variables you are to use in the problems below. The variable names are `x1`, `x2`, `p1`, `f1`. To see the value of the variable simply enter it at the prompt and type “Return”.

To submit this homework save a transcript of the *R* session in which you complete it and e-mail to me. You must include your name as a comment on the first line and in the file name.

Learning Sets are given below as tasks I want you to do to learn some material on your own. They *may* be relevant to homework problems.

Exer. 1.1. Given `x1` form a new vector `y1` whose entries are the entries of `x1` with odd index; i.e., `x[1]`, `x[3]`,

Exer. 1.2. Form a vector `y2` such that for each `j`, `y2[j]=x1[j]`, if `x1[j] > 0`, and `y2[j] = 0`, if `x1[j] ≤ 0`.

Exer. 1.3. Form a vector `y3` such that for each `j`, `y[j]=log(x2[j])`. Try the same with `x1` and report the result. Note: This is the *natural* logarithm. `log2` is the log base 2.

Exer. 1.4. Given a discrete random variable X with values X_i , for $i \leq n$, let p_i denote the probability of X_i . Let μ denote the mean of X ; i.e., $\mu = \sum_{i=1}^n X_i p_i$. The r^{th} moment about the mean of X is $\mu_r = \sum_{i=1}^n (X_i - \mu)^r p_i$. Consider `x1` as a random variable with probability distribution `p1`. Compute μ_3 of `x1`. (μ_3 is related to the *skewness* of the distribution.)

Learning Set 1.A. Vectors of one class can sometimes be coerced into a vector of another class. Look up the help entries for `as.character`, `as.numeric`, `as.integer`. Create vectors of varying types and experiment with the result of trying to coerce it to another type. (Don't forget logical vectors.)

Exer. 1.5. Create a vector `y4` whose i^{th} entry is 0 if `x1[i] > 0` and 1 if `x1[i] ≤ 0`.

Exer. 1.6. Two different students record gender of samples, one using `M`, `F` and the other using `m`, `f`. The result is in the factor `f1`. Produce a factor `f2` in which only `M`, `F` are used. Check the levels of `f2`. Can you produce such an `f2` that has only the two levels `M` and `F`?