

Clear the workspace:

```
rm(list=ls()) #clear work area
```

Specify the target subdirectory:

```
datadir="nure20141"
```

Read the 70k equally spaced normal abilities file:

```
load("Rdata/samplen70.Rdata")
ls()
## [1] "datadir" "thetas"
```

Read the names of the files in the target subdirectory into a list:

```
files=list.files(datadir) #list files in target subdirectory
n=length(files) #count of files
print(n)
## [1] 200
```

Read the files from the target directory and extract the growth scores (SGP), constructing a new dataframe with one column for the growth scores in each file:

```
for (i in 1:n){ #read each file and extract SGP values
  fname=paste(datadir,"/",files[[i]],sep="")
  load(fname) #file to load
  if(i==1){ #initialize data frame with column of lengths
    df=data.frame(seq(1,nrow(MCAS_sgp$Panel_Data)))
    colnames(df)=c("ID")
  }
  df=cbind(df,MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP) #add column for SGP values
}
str(df)

## 'data.frame': 70000 obs. of 201 variables:
##   $ ID : int 1 2 3 4 5 6 7 8 9 10 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 39 7 7 15 7 35 3 29 7 30 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 30 23 3 34 27 37 3 3 3 39 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 27 36 6 6 10 31 35 45 9 36 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 6 3 20 3 30 49 10 7 9 43 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 43 34 29 39 9 3 3 10 6 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 38 42 3 28 42 28 3 7 38 ...
##   $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 10 8 41 22 61 5 46 6 45 39 ...
```

```

## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 42 17 55 38 29 29 23 16 23 5 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 13 39 28 10 28 52 28 10 7 4 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 1 8 9 9 42 7 15 4 37 35 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 4 9 31 9 9 46 40 3 9 40 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 47 25 25 7 4 7 7 34 43 55 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 37 2 30 43 2 38 4 7 30 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 6 39 43 47 9 24 4 2 30 4 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 27 24 29 24 4 15 22 9 27 15 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 44 36 30 44 9 26 7 3 39 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 16 6 29 29 11 9 9 9 33 33 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 1 6 6 6 5 39 9 7 9 6 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 5 21 46 7 40 40 5 7 40 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 15 4 7 43 39 6 28 38 39 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 34 14 6 9 4 6 9 36 38 6 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 28 43 7 43 6 28 10 28 38 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 4 23 29 9 29 10 29 29 22 36 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 30 35 10 15 35 6 49 2 22 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 33 5 33 7 2 2 25 15 33 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 38 42 5 4 31 38 10 27 9 1 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 52 12 26 6 39 10 10 44 9 6 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 29 43 1 35 6 6 9 7 7 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 14 8 5 12 1 2 5 18 35 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 38 4 10 6 3 6 30 6 14 26 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 26 4 39 5 4 15 4 9 29 13 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 30 39 6 24 24 43 10 6 49 57 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 2 35 4 38 15 5 9 7 15 22 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 24 9 24 9 40 40 24 50 40 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 7 10 1 7 7 7 21 4 3 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 7 7 30 23 23 4 8 30 44 38 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 2 22 25 43 1 29 39 8 24 16 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 43 15 27 14 14 37 7 1 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 14 3 43 43 1 39 27 6 47 39 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 57 17 5 40 44 40 5 28 47 5 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 39 34 1 9 4 15 34 41 7 29 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 19 9 34 3 7 42 30 30 15 84 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 10 36 10 4 29 3 9 7 34 34 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 1 30 43 10 1 21 32 29 11 4 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 6 39 39 6 20 31 39 32 6 28 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 16 9 40 3 3 29 40 45 16 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 24 26 25 38 4 3 6 39 26 14 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 30 9 3 1 5 9 38 6 7 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 36 38 10 54 38 15 6 1 38 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 10 31 65 9 9 31 39 1 44 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 29 43 1 35 6 6 9 7 7 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 22 6 39 37 6 39 14 3 39 6 ...

```

```

## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 29 6 6 6 46 3 36 29 6 10 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 26 7 38 26 43 3 6 15 28 46 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 30 30 30 31 4 38 9 6 10 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 42 7 39 4 45 3 5 14 15 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 2 8 17 17 42 42 42 17 8 17 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 4 41 44 39 5 2 6 2 41 37 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 2 35 12 40 40 12 23 22 54 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 15 15 25 9 36 36 75 1 15 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 30 9 22 9 16 43 9 30 24 1 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 6 39 29 34 10 39 9 10 6 3 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 39 29 8 4 38 10 49 29 32 5 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 36 9 23 5 9 39 39 3 29 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 27 5 14 4 27 40 27 27 7 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 22 7 22 1 39 5 5 21 28 41 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 20 4 30 27 9 9 20 4 9 42 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 7 2 7 39 7 52 41 7 4 29 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 8 41 13 35 8 15 8 39 63 35 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 7 5 2 7 29 1 35 14 40 2 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 28 10 6 19 6 14 3 42 35 1 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 35 24 15 9 10 43 51 5 45 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 8 3 30 44 5 3 22 9 9 8 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 37 12 47 4 28 29 28 53 57 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 41 41 31 9 49 9 27 1 41 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 6 42 6 34 43 38 42 9 3 42 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 48 40 40 23 28 5 15 32 5 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 6 38 7 6 38 43 43 6 10 26 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 44 9 41 41 2 4 7 4 15 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 39 44 9 30 39 21 6 14 5 30 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 26 19 30 30 39 17 44 3 5 5 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 42 39 42 9 5 14 46 28 9 40 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 56 7 6 14 6 3 7 3 14 38 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 46 2 29 35 7 29 2 7 9 29 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 16 4 5 40 4 5 44 37 35 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 3 34 15 14 9 3 8 42 15 6 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 22 4 38 34 39 3 38 34 22 4 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 20 30 18 1 3 37 5 30 7 15 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 29 8 15 33 28 21 8 29 34 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 29 3 43 39 31 29 33 21 39 21 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 9 38 9 9 16 33 39 3 9 3 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 8 42 8 1 4 42 34 37 42 39 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 21 4 8 3 38 36 36 38 7 40 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 27 3 27 51 5 3 56 25 7 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 7 9 42 42 6 5 7 31 24 1 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 3 36 3 10 9 37 5 28 9 ...
## $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int 5 9 9 5 41 30 50 1 8 55 ...

```

```
##  $ MCAS_sgp$SGPercentiles$MATHEMATICS.2010$SGP: int  5 39 9 9 28 37 5 32 5 70 ...
##  [list output truncated]
```

Now we have a data frame with a row for each "student" (theta value) and a column for each replication (file).

R allows us to easily compute a new column that is the mean of the SGP scores in a row (i.e., the mean SGP for that student over all of the replications). This produces a vector of 70,000 means, one for each "student":

```
means=apply(df[,2:ncol(df)],1,mean) #compute vector of row means exclude first column
```

We can do the same for the standard deviation of the SGP scores:

```
sds=apply(df[,2:ncol(df)],1,sd) #compute vector of row standard deviations exclude first column
```

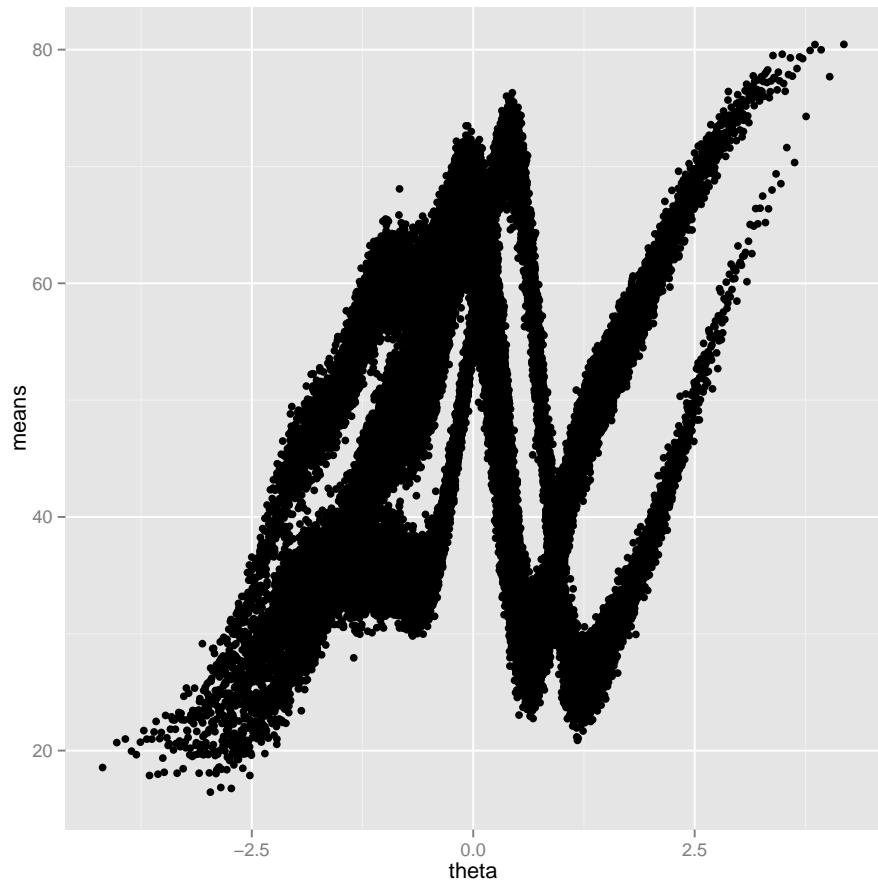
Now we want to plot the results with ggplot2. Ggplot2 requires you to supply the data to be plotted in the form of a data frame, so we have to build one.

We'll put the ability parameter  $\theta$  on the horizontal axis. Our data frame has to include any variables we use in the plot, so we have to have the thetas as well as the means and standard deviations:

```
theta=thetas$theta
gdf=data.frame(theta,means,sds)
```

Now that we have the data in the form ggplot2 requires, we can plot the means as a function of theta:

```
library(ggplot2)
p=ggplot(gdf,aes(x=theta,y=means))
p+geom_point()
```



Next, plot the standard deviations as a function of theta:

```
library(ggplot2)
p=ggplot(gdf,aes(x=theta,y=sds))
p+geom_point()
```

