Testing Times

Ralf Martin

# Dice

How likely is it to throw repeated sixes with a fair dice? Let’s use a loop to work it out

for(i in 1:5){  
 p=as.character(format( round(1/6^i,4),scientific=F))  
   
 cat(paste0("$$Prob\\{", i,"\\textrm{ sixes in a row}\\}=",p,"$$\n"))  
 # Cat is like print but better suited for use in markdown where mingle  
 # other stuff into the output such as latex  
}

# When should we start taking the slope of a line serious?

library(dplyr)

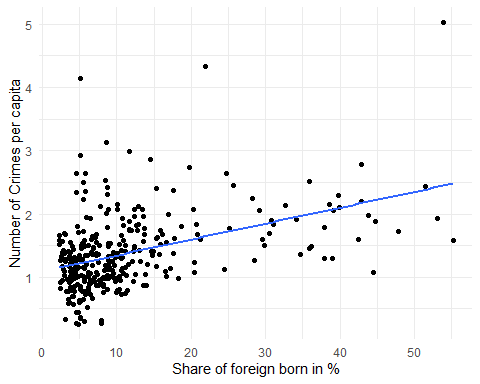
##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)  
ff=read.csv("https://www.dropbox.com/s/g1w75gkw7g91zef/foreigners.csv?dl=1")   
ff=ff%>%mutate(crimesPc=crimes11/pop11)%>% filter( crimesPc<10) # We are dropping some outliers  
  
ggplot(ff, aes(y=crimesPc,x=b\_migr11)) + geom\_point() + # Make sure + is on this line so R understands the command is not finished  
 xlab("Share of foreign born in %") +   
 ylab("Number of Crimes per capita") +   
 geom\_smooth(method = "lm", se = FALSE) +  
 theme\_minimal()

## `geom\_smooth()` using formula 'y ~ x'



Recall the regression command:

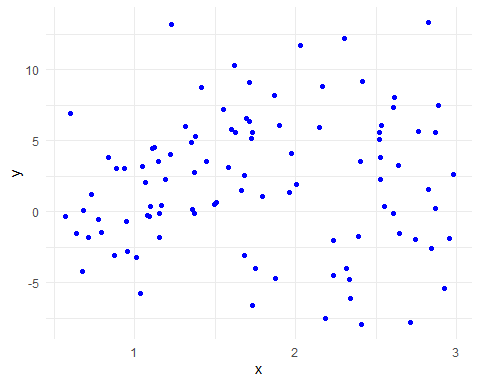
df=ff  
reg1=lm(crimesPc~b\_migr11,df)   
reg1 %>% summary()

##   
## Call:  
## lm(formula = crimesPc ~ b\_migr11, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.13314 -0.33959 -0.06763 0.22302 2.92572   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.091273 0.045146 24.17 < 2e-16 \*\*\*  
## b\_migr11 0.025164 0.002922 8.61 3.33e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5482 on 321 degrees of freedom  
## Multiple R-squared: 0.1876, Adjusted R-squared: 0.1851   
## F-statistic: 74.14 on 1 and 321 DF, p-value: 3.325e-16

# Monte Carlo

Let’s create the data ourselves so we know what drives it.

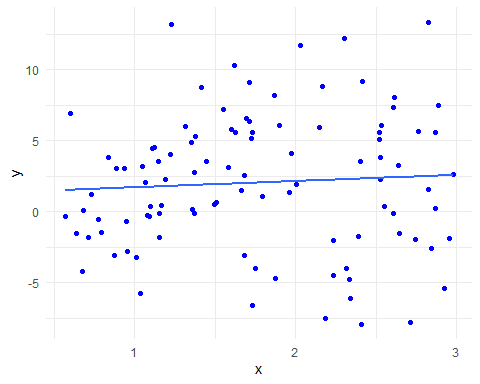
obs <- 100  
 x <- 0.5 + runif(obs)\*2.5  
 sig=sqrt(5.5)\*2  
 eps <- rnorm(obs,0,sig)  
 y <- 2 + x \* 0 + eps  
   
 df=data.frame(x,y)  
   
 ggplot(df, aes(x, y))+geom\_point(color="blue") +theme\_minimal()



Now let’s run a regression

ggplot(df, aes(x, y))+geom\_point(color="blue") +theme\_minimal()+geom\_smooth(method="lm",se=F)

## `geom\_smooth()` using formula 'y ~ x'



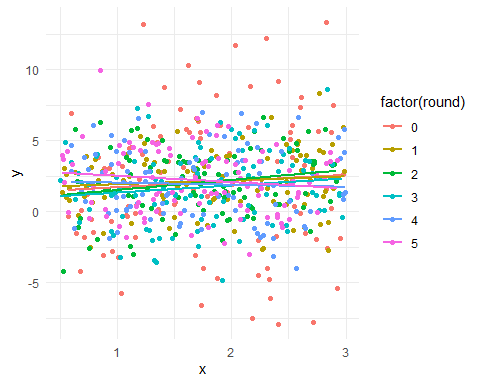
monte1 <- lm(y ~ x , data = df)  
   
 summary(monte1)

##   
## Call:  
## lm(formula = y ~ x, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.251 -3.138 -0.009 3.272 11.373   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1.2551 1.3032 0.963 0.338  
## x 0.4345 0.6874 0.632 0.529  
##   
## Residual standard error: 4.794 on 98 degrees of freedom  
## Multiple R-squared: 0.004061, Adjusted R-squared: -0.006102   
## F-statistic: 0.3996 on 1 and 98 DF, p-value: 0.5288

Now let’s repeat that many times

df=df %>% mutate(round=0)  
 b2<-numeric()  
   
 sig=sqrt(5.5)  
   
 for(i in 1:5) {  
   
 x <- 0.5 + runif(obs)\*2.5   
   
 eps <- rnorm(obs,0,sig)  
 y <- 2 + x \* 0 + eps  
   
 monte2 <- lm(y ~ x )  
 b2[i]=monte2$coefficients[2]  
   
   
 bx =monte2$coefficients[2]   
   
 dfnew=data.frame(x,y,round=i)  
 df=bind\_rows(df,dfnew)  
 }  
   
   
 ggplot(df, aes(x, y,color=factor(round)))+geom\_point() +theme\_minimal()+geom\_smooth(method="lm",se=F)

## `geom\_smooth()` using formula 'y ~ x'



Really a lot of times..

df=df %>% mutate(round=0)  
 b2<-numeric()  
   
 sig=sqrt(5.5)  
   
 for(i in 1:1000) {  
   
 x <- 0.5 + runif(obs)\*2.5   
   
 eps <- rnorm(obs,0,sig)  
 y <- 2 + x \* 0 + eps  
   
 monte2 <- lm(y ~ x )  
 b2[i]=monte2$coefficients[2]  
   
   
 bx =monte2$coefficients[2]   
   
 dfnew=data.frame(x,y,round=i)  
 df=bind\_rows(df,dfnew)  
 }   
   
#b2[i]=monte2$coefficients[2]   
bx

## x   
## 0.2529307

b2

## [1] -2.089435e-01 -5.157838e-01 7.917320e-03 1.860723e-01 -4.883173e-01  
## [6] -2.028619e-01 -4.371328e-01 -8.574932e-01 9.103888e-02 7.171228e-01  
## [11] 3.126729e-01 -1.671843e-01 1.939506e-01 -8.250124e-02 6.024176e-01  
## [16] -4.255966e-01 3.267966e-01 7.574939e-02 -1.830506e-02 -2.660495e-01  
## [21] 1.373053e-01 -3.899971e-01 3.544639e-01 -1.434018e-01 1.053305e-02  
## [26] -6.135168e-01 4.948900e-01 -6.481225e-01 1.730771e-01 -2.752519e-01  
## [31] -2.100278e-01 -3.678867e-01 -2.080212e-01 -8.002357e-02 -1.362385e-01  
## [36] -1.707988e-01 2.635435e-01 -1.879758e-01 2.046394e-01 7.360455e-02  
## [41] -1.002451e-02 -5.488155e-02 -1.867020e-02 3.016584e-01 -2.368462e-01  
## [46] -2.579833e-01 4.693119e-01 3.403458e-01 -2.778187e-01 -2.747383e-01  
## [51] 3.214378e-01 1.466242e-01 2.664213e-01 -4.668630e-02 -1.126723e-01  
## [56] -1.898007e-01 1.113674e-01 5.073378e-01 1.882520e-01 -4.707656e-01  
## [61] 3.414882e-01 3.793250e-01 2.807336e-01 -1.657438e-01 -1.073944e-01  
## [66] -6.042826e-01 -1.005437e-01 2.412092e-02 2.829938e-01 3.540662e-01  
## [71] -5.200686e-01 -5.150093e-02 -1.725681e-01 -2.977563e-01 2.038488e-01  
## [76] -6.955881e-01 1.675198e-01 -5.071196e-01 3.470279e-01 -6.811426e-01  
## [81] -5.228195e-01 4.726321e-01 2.715836e-01 -1.027434e-01 -7.647946e-02  
## [86] 2.980753e-01 1.805852e-01 6.270123e-02 -1.667827e-01 1.390673e-01  
## [91] -1.176840e-02 -7.823193e-01 6.021267e-01 -1.232170e-01 -2.768872e-01  
## [96] 3.613472e-02 3.827881e-01 -1.380888e-01 -1.381925e-01 1.987429e-01  
## [101] -2.038693e-01 -3.005726e-01 -1.737585e-02 2.179189e-01 1.366212e-01  
## [106] -4.157156e-01 4.814712e-02 1.596615e-01 1.494433e-01 -1.671473e-01  
## [111] 1.002775e-01 -6.056475e-01 -2.006435e-01 -4.672733e-01 4.253790e-01  
## [116] 6.423761e-01 -1.348369e-01 -1.322848e-01 2.176302e-01 1.339260e-01  
## [121] -1.080720e-01 -5.545253e-02 -2.381168e-01 5.736855e-02 -4.428716e-02  
## [126] -4.454686e-02 2.483964e-01 6.374575e-01 3.196758e-01 -3.788995e-01  
## [131] -2.022869e-01 7.952308e-02 5.075639e-02 4.819926e-01 1.278090e-01  
## [136] -2.324468e-01 -6.495052e-01 3.986024e-01 -1.680039e-01 1.911286e-01  
## [141] -1.291780e-01 -3.448986e-01 -4.238081e-01 -3.084510e-02 1.243014e-01  
## [146] 5.367584e-05 1.793744e-01 1.248529e-01 -2.414354e-01 -6.344812e-02  
## [151] -7.039808e-02 7.407404e-01 1.220975e-01 -7.217130e-01 7.170755e-03  
## [156] 3.505817e-01 -2.579733e-01 3.064334e-01 -5.380284e-01 -2.227280e-01  
## [161] -1.701271e-01 7.111309e-02 3.660995e-01 -3.830851e-01 1.448848e-01  
## [166] -3.394136e-01 1.745428e-02 -5.733802e-01 2.145259e-01 -1.204484e-01  
## [171] -3.831345e-01 2.504745e-02 1.235568e-02 -1.601765e-03 -9.682182e-02  
## [176] -2.584859e-01 -5.730811e-02 3.425673e-02 4.265186e-02 -2.655564e-01  
## [181] -5.100148e-01 -4.405079e-01 1.930115e-01 3.939817e-01 2.636498e-01  
## [186] -5.064175e-01 7.804167e-01 -3.889061e-01 6.173157e-01 1.888245e-01  
## [191] -8.001212e-02 -1.717722e-01 1.880914e-02 6.053610e-01 2.091225e-01  
## [196] -4.316624e-01 1.144835e-01 1.983617e-02 6.166096e-01 -4.662269e-02  
## [201] -4.983134e-01 4.537895e-01 -1.364762e-01 2.186539e-01 5.224157e-01  
## [206] 2.717214e-01 4.310302e-03 4.593308e-01 -3.268801e-01 1.729506e-01  
## [211] 5.900557e-01 4.133971e-01 5.593211e-01 -8.744685e-01 3.820851e-01  
## [216] -4.585792e-01 -2.611923e-01 5.794692e-01 -9.154295e-02 3.371816e-01  
## [221] -2.953782e-01 -5.165365e-01 -5.282056e-02 -3.915439e-01 -4.294381e-01  
## [226] -3.917531e-01 6.054074e-03 -2.735250e-01 6.548014e-02 1.178847e-01  
## [231] -4.233934e-01 2.298300e-01 -3.046293e-01 5.998458e-01 3.142121e-01  
## [236] 1.074740e-02 -3.183846e-01 8.193332e-02 -2.743668e-02 3.116852e-02  
## [241] -5.597772e-01 -2.955089e-01 -2.360733e-02 -5.811192e-02 -5.330518e-01  
## [246] 2.081192e-02 3.469615e-01 -1.234158e-01 -4.477286e-01 -1.968291e-01  
## [251] -2.432337e-02 -2.495600e-03 -6.640385e-03 3.546227e-02 3.983967e-01  
## [256] 8.996448e-01 2.666940e-01 -2.652819e-01 -3.680823e-01 2.417027e-01  
## [261] 9.501436e-02 3.777376e-01 -4.067208e-01 -8.031783e-01 -4.261908e-01  
## [266] 1.710950e-01 -3.139699e-01 8.565062e-01 1.874470e-01 -2.562915e-01  
## [271] -3.772724e-02 3.608174e-01 2.723194e-01 -3.701048e-01 -2.921603e-01  
## [276] 4.226314e-01 4.610769e-02 2.278001e-01 -5.840532e-01 2.679312e-01  
## [281] 5.821355e-01 2.882161e-01 6.592981e-02 -2.021001e-01 6.219832e-02  
## [286] 6.936090e-01 -2.764894e-01 1.528059e-01 -1.727641e-02 1.109687e-01  
## [291] -6.979928e-03 3.889304e-01 5.862265e-02 5.370186e-01 -1.136329e-02  
## [296] -4.113342e-01 -3.713643e-02 2.114293e-01 -9.280147e-01 -1.275917e-01  
## [301] 3.758649e-01 5.603399e-02 3.201071e-01 -3.343893e-01 -1.797979e-01  
## [306] 4.179372e-01 -1.784968e-01 -7.596602e-01 -4.201632e-01 -8.552707e-02  
## [311] -1.817453e-01 -3.630850e-01 -1.857040e-01 2.981077e-01 1.171006e-01  
## [316] 3.048046e-01 -6.386265e-01 -5.886455e-01 2.448075e-01 -1.181301e-01  
## [321] 6.197242e-02 3.782610e-01 -6.400692e-01 -6.865150e-02 4.704599e-01  
## [326] -4.629901e-01 -1.328029e-01 -4.679723e-01 -2.644662e-01 4.312074e-02  
## [331] -5.288513e-03 -3.529410e-01 5.459623e-01 2.089347e-02 -1.890880e-01  
## [336] 1.868310e-01 -3.751389e-01 2.465161e-03 5.154402e-01 1.863223e-01  
## [341] 4.474840e-01 -3.705604e-01 -1.277508e-01 3.055632e-01 8.519299e-01  
## [346] -2.524198e-01 1.112512e-01 3.850628e-01 5.337761e-01 -2.440361e-01  
## [351] 3.075876e-01 2.148855e-01 5.110815e-01 1.553336e-01 -2.843842e-01  
## [356] -4.915067e-01 -1.860078e-03 -1.901442e-01 -3.849845e-01 6.711075e-01  
## [361] 2.626973e-01 4.021152e-01 2.367065e-01 -2.085185e-01 -4.164771e-01  
## [366] -3.247021e-01 -5.571863e-01 6.535450e-01 -1.555735e-02 5.241410e-01  
## [371] -4.283691e-01 -4.602756e-01 -2.090077e-01 1.698474e-01 1.530172e-01  
## [376] 3.559503e-01 5.794913e-01 -4.638128e-01 -2.578089e-01 2.302854e-01  
## [381] 7.861257e-02 4.584161e-02 -3.407148e-01 3.901303e-01 -1.324785e-01  
## [386] 2.637810e-01 -3.044611e-01 -2.248637e-01 -5.217397e-02 -6.912134e-02  
## [391] -1.274407e-02 -6.071791e-02 4.597838e-02 -2.655911e-01 -5.527408e-01  
## [396] -4.303686e-01 2.827325e-01 1.491858e-01 2.008902e-03 -6.864942e-02  
## [401] -1.748808e-01 4.229123e-01 5.488372e-03 -7.864820e-02 2.350374e-02  
## [406] -5.308540e-01 -1.244715e-01 4.135101e-01 -2.007588e-01 2.734169e-02  
## [411] 1.167680e-02 2.625927e-01 1.305050e-01 -2.377425e-02 -3.110856e-01  
## [416] -6.732058e-02 4.872957e-02 -2.936892e-01 5.376436e-01 7.248740e-01  
## [421] -4.563716e-01 5.593572e-01 9.379210e-02 -3.828066e-01 4.035682e-01  
## [426] 4.396231e-01 6.878711e-01 8.911960e-02 2.183727e-02 -8.811791e-02  
## [431] 3.017599e-01 -1.224461e-01 -2.214009e-01 3.968434e-01 -2.614860e-01  
## [436] 4.486424e-01 -2.271662e-01 4.402518e-01 2.012580e-01 1.992097e-01  
## [441] 3.514085e-01 4.666335e-01 8.602424e-01 3.022848e-01 -5.410160e-01  
## [446] 3.086785e-01 4.880935e-01 3.077359e-01 -3.854484e-01 2.822227e-01  
## [451] -2.651336e-01 2.812069e-03 -2.979871e-01 4.395205e-01 -2.519058e-01  
## [456] -2.284543e-01 3.476567e-01 -2.566419e-01 1.554807e-01 -3.421251e-01  
## [461] -4.038047e-01 -4.287405e-02 -1.683536e-01 -4.778455e-01 -6.708972e-01  
## [466] -8.274997e-02 -4.553371e-01 -3.441108e-02 -2.945158e-01 2.369604e-01  
## [471] 2.095733e-01 -6.175163e-02 6.594777e-02 -3.161312e-01 -1.795603e-01  
## [476] -3.608211e-01 -3.700447e-01 -3.487435e-01 -8.262834e-02 -1.590543e-01  
## [481] 1.753859e-01 4.320447e-01 4.618072e-01 -8.720228e-02 9.962450e-02  
## [486] 4.459303e-02 -1.493966e-01 -3.211296e-02 7.312001e-02 3.253098e-01  
## [491] -3.333184e-01 -5.083692e-01 -3.878390e-01 3.088437e-01 4.952994e-02  
## [496] -2.026875e-01 -7.441932e-01 -2.175785e-01 -2.661521e-03 2.363842e-01  
## [501] -1.133572e-01 -4.352962e-01 -4.072840e-02 6.126699e-01 -7.900885e-03  
## [506] -4.589789e-03 -2.193363e-02 -3.663562e-01 1.804071e-01 -2.805810e-02  
## [511] -5.338233e-01 -1.870321e-01 -3.503262e-01 -9.811021e-02 -3.808283e-01  
## [516] -2.884641e-01 4.643555e-01 1.900197e-01 -5.639849e-01 -1.610363e-01  
## [521] 3.081901e-01 -3.629229e-02 -2.382852e-02 3.403617e-02 -1.282046e-01  
## [526] -5.601346e-03 1.241952e-01 -5.056116e-01 1.441337e-02 -3.101114e-01  
## [531] 3.850862e-02 1.653062e-01 -5.885270e-01 -5.286667e-01 4.297668e-01  
## [536] -3.885100e-01 -1.537423e-01 7.953826e-02 1.955794e-01 -1.823439e-01  
## [541] 6.458964e-02 -2.997481e-01 -2.906379e-01 -1.216327e-03 -2.839276e-01  
## [546] 1.457583e-01 -2.148599e-01 7.832837e-02 -2.346074e-01 -3.243541e-01  
## [551] 7.794377e-02 -2.445061e-01 5.169835e-02 2.329396e-01 -2.329692e-01  
## [556] 3.365165e-01 -6.456475e-02 -9.672273e-02 5.025527e-02 3.311639e-01  
## [561] -4.273126e-01 -2.310715e-01 4.644505e-02 1.272128e-01 6.582102e-01  
## [566] 1.031585e-03 -3.220965e-03 3.686352e-01 5.998895e-01 4.062090e-01  
## [571] 6.210016e-01 3.587239e-02 1.636751e-01 2.610254e-01 9.539106e-02  
## [576] 4.765527e-01 2.537049e-01 -1.721201e-01 -1.892763e-02 -4.714559e-01  
## [581] 4.217925e-01 -4.896257e-01 -3.538637e-01 2.820608e-01 -1.101776e-01  
## [586] -8.900258e-01 2.909762e-01 3.516965e-01 -5.658644e-01 -4.092099e-01  
## [591] -7.745393e-02 3.182112e-01 -3.758973e-02 2.904167e-01 -5.188657e-01  
## [596] -6.035699e-01 4.338927e-02 -2.281405e-01 -8.186020e-02 3.894852e-01  
## [601] 5.871138e-02 -6.760139e-01 2.155523e-01 -4.826117e-01 -2.030630e-01  
## [606] -1.915087e-01 -1.155211e-01 -1.068067e-01 3.451281e-01 -5.420040e-01  
## [611] -2.839483e-01 -3.532472e-01 3.821400e-02 -2.843479e-01 1.009616e-01  
## [616] -1.271430e-01 -3.958402e-01 3.419877e-02 1.291179e-01 -3.431704e-02  
## [621] -2.317410e-01 9.308764e-03 5.505971e-01 -6.235114e-02 5.330263e-01  
## [626] -3.498162e-02 5.423964e-02 -4.423266e-02 3.629089e-02 -1.044742e-01  
## [631] -1.474821e-01 7.414315e-02 8.603367e-02 -1.009346e-01 1.651726e-01  
## [636] -1.828988e-01 6.377576e-01 2.707606e-02 2.498038e-01 3.265890e-01  
## [641] -7.619428e-01 1.776702e-01 -6.120634e-01 -6.476993e-01 -4.520673e-02  
## [646] 3.241751e-01 -2.649479e-02 -5.292463e-01 -2.970082e-01 -2.839228e-01  
## [651] -6.071346e-01 -2.908523e-01 -1.230818e-01 -2.471729e-01 4.151374e-01  
## [656] 1.884377e-01 -2.664341e-01 -2.345213e-02 -1.862726e-01 2.710882e-01  
## [661] -5.177175e-01 -2.028556e-01 7.511998e-01 -5.906966e-01 7.974215e-01  
## [666] -5.915114e-01 -1.916651e-01 -7.965864e-02 -3.054919e-01 8.865095e-02  
## [671] -3.764185e-01 -4.575839e-01 -2.361298e-01 8.320621e-02 -1.445369e-01  
## [676] 2.956643e-01 -1.693158e-01 1.782335e-01 -2.543750e-01 -2.787335e-01  
## [681] -2.651083e-02 2.182458e-01 3.118868e-01 -6.256680e-02 7.458300e-02  
## [686] 2.146791e-01 -1.106769e-01 -3.423296e-01 9.272522e-02 -2.071473e-01  
## [691] -1.780207e-02 -3.298653e-01 -4.060660e-01 -2.721640e-01 3.514847e-01  
## [696] -2.592284e-01 2.290773e-01 3.221521e-01 -4.698099e-01 -4.320062e-01  
## [701] 3.865506e-01 -1.736426e-01 -8.887117e-04 -3.459436e-02 -6.560801e-01  
## [706] 3.657775e-02 -1.258273e-01 2.485070e-03 -5.169281e-01 3.216308e-01  
## [711] -7.409801e-01 -7.362216e-02 -2.968794e-01 -4.776843e-01 7.866728e-01  
## [716] -9.204614e-02 4.713171e-01 2.670801e-01 4.733580e-01 1.559942e-01  
## [721] 3.592556e-01 -5.632484e-02 1.733643e-02 -2.706472e-01 -1.918266e-01  
## [726] -3.668724e-01 1.593613e-01 -2.547149e-01 9.818667e-02 -4.964902e-01  
## [731] -3.180406e-02 -3.364075e-01 3.064209e-01 2.770665e-01 4.139216e-01  
## [736] -3.990689e-01 3.255127e-01 2.022985e-01 -1.425602e-01 -6.464696e-01  
## [741] 5.342989e-01 3.613795e-02 7.355857e-02 -9.938734e-01 5.752083e-02  
## [746] -5.326802e-01 5.155349e-02 8.744390e-02 -1.943860e-01 3.596433e-01  
## [751] -2.701088e-01 2.972215e-01 -3.819496e-01 3.136238e-01 -5.792723e-02  
## [756] 1.200793e-01 3.456504e-01 -3.515115e-01 1.424279e-01 1.529828e-01  
## [761] -5.763512e-02 -1.984606e-01 9.205512e-02 -3.425075e-01 -1.577414e-01  
## [766] 3.259781e-01 -1.728785e-01 3.786272e-01 6.229605e-01 -1.086392e-01  
## [771] -4.311319e-01 3.564919e-01 -2.468328e-01 5.324703e-01 1.774593e-01  
## [776] -2.013246e-01 3.138657e-01 -4.085716e-01 1.053751e-01 -5.995023e-01  
## [781] -1.616309e-01 -1.751171e-02 -2.584535e-01 -2.978048e-01 -5.596832e-01  
## [786] 5.399541e-01 6.164294e-02 -7.844979e-02 -4.176275e-01 2.347878e-01  
## [791] 6.627000e-01 5.374809e-02 -1.885458e-02 5.208678e-01 -5.851443e-02  
## [796] 3.107760e-01 -4.350035e-02 -7.441166e-02 -6.444171e-01 -1.056799e-01  
## [801] 1.854523e-01 1.012162e+00 -2.151011e-01 4.982472e-01 -5.360925e-01  
## [806] 4.749773e-01 9.905102e-02 3.153533e-01 1.406659e-01 -4.639424e-01  
## [811] -3.536923e-01 -2.223759e-01 -5.468942e-01 -3.796613e-01 3.314483e-02  
## [816] 5.543923e-02 7.723989e-02 5.408480e-01 -2.852410e-01 8.175429e-01  
## [821] -8.980060e-02 -1.871198e-01 -1.132240e-01 5.240615e-01 1.948309e-01  
## [826] 1.258871e-01 -2.521954e-01 -3.529105e-01 4.712603e-02 -5.478272e-02  
## [831] -3.341470e-01 3.773041e-01 -3.047720e-02 3.731940e-02 -4.334017e-01  
## [836] -1.338518e-01 -6.197458e-01 -2.021280e-01 -4.981550e-01 2.874783e-01  
## [841] -3.153394e-01 -2.077389e-01 9.990958e-02 3.525472e-01 -2.135285e-01  
## [846] -2.503348e-01 -2.807646e-01 1.002117e-01 3.512913e-01 -4.940173e-02  
## [851] 2.341665e-01 4.132178e-01 -1.590772e-01 1.758676e-01 -3.048310e-02  
## [856] 2.844381e-01 -6.589861e-01 -1.625829e-01 -1.265893e-01 2.745366e-01  
## [861] -4.398893e-01 1.387478e-01 4.569160e-02 -6.798774e-01 -2.186145e-02  
## [866] 3.674263e-02 3.619607e-01 -1.152173e-01 2.589544e-01 -2.440954e-01  
## [871] 2.406353e-01 -2.964152e-01 6.533135e-02 2.888572e-02 3.004103e-01  
## [876] -9.766938e-02 -2.277441e-01 1.131729e-01 -3.461982e-01 1.954672e-02  
## [881] 2.062966e-01 4.395985e-01 -2.923941e-04 -4.586839e-01 5.404639e-01  
## [886] 1.116757e-01 1.852137e-01 -2.667741e-01 1.710272e-01 -4.823348e-01  
## [891] -1.226106e-01 7.614604e-01 -3.636338e-02 1.798577e-01 1.662538e-01  
## [896] 2.985882e-01 -5.145934e-01 2.121146e-01 -4.809469e-01 -2.159566e-01  
## [901] -1.258583e-01 2.010404e-02 8.499487e-01 5.820417e-02 -4.343812e-01  
## [906] 3.300693e-01 -3.227969e-01 -2.293454e-02 1.402444e-01 2.286119e-01  
## [911] -7.325303e-02 1.339629e-01 4.836105e-02 2.481883e-01 2.429119e-01  
## [916] 3.488636e-01 -3.049494e-02 -7.631338e-02 -3.363543e-01 6.753431e-01  
## [921] -7.253377e-01 2.692960e-01 1.891838e-01 1.765793e-01 4.972487e-01  
## [926] 7.427969e-02 -3.823026e-01 -6.269491e-02 -7.367799e-02 2.015837e-01  
## [931] 1.701565e-01 5.156174e-01 6.346693e-02 6.692009e-02 4.218748e-01  
## [936] -7.023971e-01 3.875487e-01 5.169836e-01 -5.137179e-01 4.158602e-01  
## [941] 3.903620e-01 -1.120259e-01 6.840650e-01 2.175136e-01 5.096894e-01  
## [946] 3.644894e-01 5.423399e-01 2.559438e-01 -5.918504e-02 4.857723e-01  
## [951] -4.729280e-01 3.782996e-01 -8.082339e-01 2.166378e-02 2.233466e-01  
## [956] -2.992633e-01 2.678318e-01 -1.324205e-01 -6.305444e-02 3.881350e-01  
## [961] 1.043467e-01 2.350819e-01 -4.623397e-01 6.771013e-01 -7.973085e-02  
## [966] -9.252943e-02 -6.381432e-02 -1.506337e-01 1.075652e-01 -3.391208e-01  
## [971] 2.477516e-01 -1.558939e-01 -5.403660e-02 -4.264700e-01 1.599269e-01  
## [976] -3.821927e-01 5.078610e-01 2.914315e-01 -1.271465e-02 6.674362e-02  
## [981] 4.441526e-02 -1.076396e-01 -1.666173e-01 1.043080e+00 -1.308417e-01  
## [986] -1.364659e-01 3.073043e-01 -2.262339e-01 1.705879e-01 1.836268e-01  
## [991] -8.630535e-02 3.129679e-01 1.341333e-01 6.911798e-01 3.922008e-01  
## [996] 6.079664e-02 1.748235e-02 7.155887e-02 2.848417e-01 2.529307e-01

b2[10000]

## [1] NA

bx

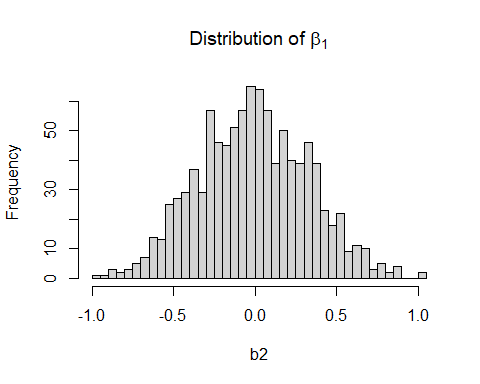
## x   
## 0.2529307

mean(b2)

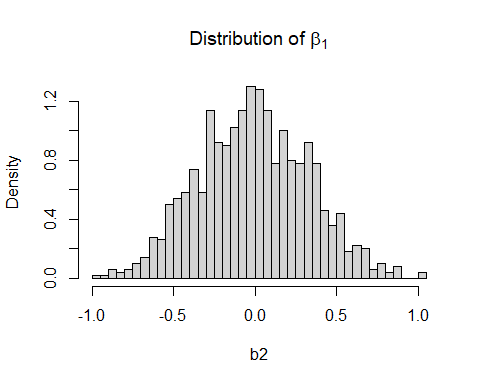
## [1] -0.009677026

Let’s look at the distribution

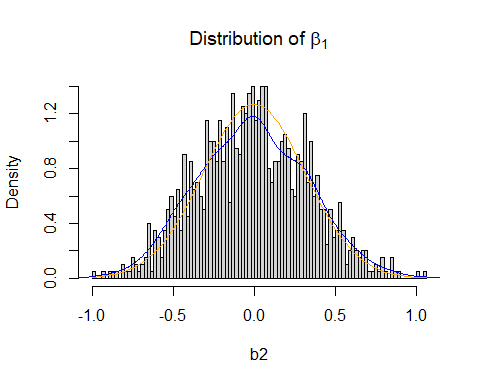
library(latex2exp)  
 # Compute the standard errors of the estimates  
 b2\_sig =sqrt(sig^2/(obs\* var(x)) )  
  
   
 hist(b2,30,main=TeX("Distribution of $\\beta\_1$"))



hist(b2,30,prob=TRUE,main=TeX("Distribution of $\\beta\_1$"))



hist(b2,100,prob=TRUE,main=TeX("Distribution of $\\beta\_1$"))  
 db2=density(b2)  
 lines(db2,col="blue")  
   
 curve(dnorm(x,0,b2\_sig),add=TRUE, col="orange")

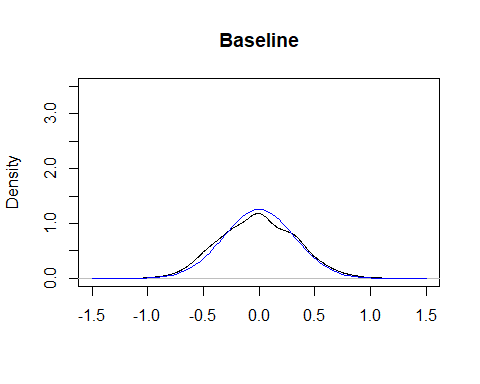


Now let’s look what happens if we have - less observations - more variation in X - more variation in epsilon - non normal epsilon - non normal epsilon less observations

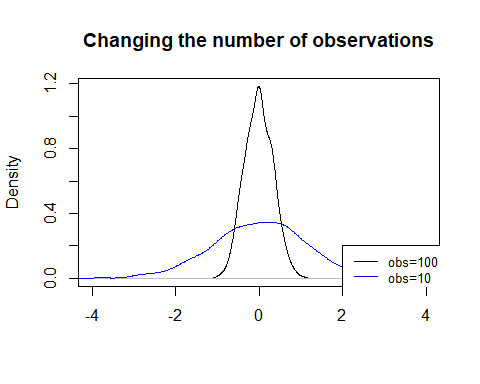
b2\_less\_obs<-numeric()  
 b2\_more\_obs<-numeric()  
 b2\_more\_var\_x<-numeric()  
 b2\_more\_var\_eps<-numeric()  
 b2\_non\_normal\_eps<-numeric()  
 b2\_non\_normal\_eps\_less<-numeric()  
 b2\_non\_normal\_eps\_more<-numeric()  
   
   
 less\_obs=10  
   
 more\_obs=750  
   
   
   
  
   
 for(i in 1:1000) {  
  
 x <- 0.5 + runif(obs) \*2.5  
 x\_less\_obs <- 0.5 + runif(less\_obs)\*2.5  
 x\_more\_var\_x <- 0.5 + runif(obs) \*5  
 x\_more\_obs <- 0.5 + runif(more\_obs)\*2.5  
  
   
  
 eps = rnorm(obs,0,sig)  
 eps\_less\_obs = rnorm(less\_obs,0,sig)  
 eps\_more\_obs = rnorm(more\_obs,0,sig)  
   
 eps\_more\_var\_eps = rnorm(obs,0,sig\*4)  
   
   
 uni=runif(obs)  
 eps\_non\_normal\_eps = -1\*(uni<0.5) + 10\*(uni>=.95)  
   
  
 uni=runif(more\_obs)  
 eps\_non\_normal\_eps\_more = -1\*(uni<0.5) + 10\*(uni>=.95)  
   
   
 beta1=0   
   
 y = 2 + x \* beta1 + eps  
 y\_less\_obs = 2 + x\_less\_obs \* beta1 + eps\_less\_obs  
 y\_more\_obs = 2 + x\_more\_obs \* beta1 + eps\_more\_obs  
   
   
 y\_more\_var\_x = 2 + x\_more\_var\_x \* beta1 + eps  
 y\_more\_var\_eps = 2 + x \* beta1 + eps\_more\_var\_eps  
   
 y\_non\_normal\_eps = 2 + x \* beta1 + eps\_non\_normal\_eps  
   
   
 y\_non\_normal\_eps\_less = 2 + x\_less\_obs \* beta1 + eps\_non\_normal\_eps[1:less\_obs]  
 y\_non\_normal\_eps\_more = 2 + x\_more\_obs \* beta1 + eps\_non\_normal\_eps\_more  
   
   
   
   
 monte2 <- lm(y\_less\_obs ~ x\_less\_obs )  
 b2\_less\_obs[i]=monte2$coefficients[2]  
   
 monte2 <- lm(y\_more\_obs ~ x\_more\_obs )  
 b2\_more\_obs[i]=monte2$coefficients[2]  
  
   
 monte2 <- lm(y\_more\_var\_x ~ x\_more\_var\_x )  
 b2\_more\_var\_x[i]=monte2$coefficients[2]  
  
   
 monte2 <- lm(y\_more\_var\_eps ~ x )  
 b2\_more\_var\_eps[i]=monte2$coefficients[2]  
   
   
 monte2 <- lm(y\_non\_normal\_eps ~ x )  
 b2\_non\_normal\_eps[i]=monte2$coefficients[2]  
   
  
 monte2 <- lm(y\_non\_normal\_eps\_less ~ x\_less\_obs )  
 b2\_non\_normal\_eps\_less[i]=monte2$coefficients[2]  
  
 monte2 <- lm(y\_non\_normal\_eps\_more ~ x\_more\_obs )  
 b2\_non\_normal\_eps\_more[i]=monte2$coefficients[2]  
   
   
 }

Let’s look at this

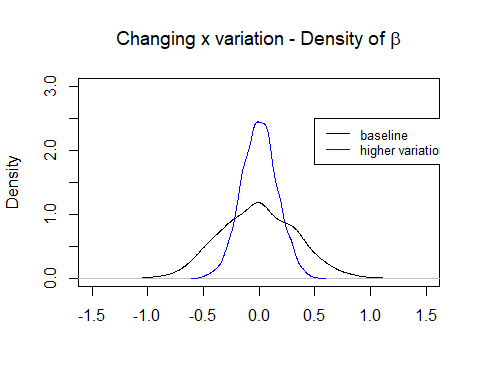
# Compute the standard errors of the estimates  
 b2\_sig =sqrt(sig^2/(obs\* var(x)) )  
  
  
 #b2\_sig\_less\_obs=sqrt(sig^2/(less\_obs\* var(x\_less\_obs)))  
 #b2\_sig\_less\_obs =sqrt(sig^2/(more\_obs\* var(x\_more\_obs)) )  
 #b2\_sig\_more\_var\_eps=sqrt((sig\*4)^2/(obs\* var(x)))   
  
  
 # Compute the densities  
 db2=density(b2)  
 db2\_less\_obs=density(b2\_less\_obs)  
 db2\_more\_obs=density(b2\_more\_obs)  
   
 db2\_more\_var\_x=density(b2\_more\_var\_x)  
 db2\_more\_var\_eps=density(b2\_more\_var\_eps)  
 db2\_non\_normal\_eps=density(b2\_non\_normal\_eps)  
 db2\_non\_normal\_eps\_less=density(b2\_non\_normal\_eps\_less)  
 db2\_non\_normal\_eps\_more=density(b2\_non\_normal\_eps\_more)  
   
 db2\_eps\_more\_var\_eps=density(eps\_more\_var\_eps)  
 db2\_eps=density(eps)  
   
 # Baseline  
 plot(db2,xlim=c(-1.5,1.5),ylim=c(0,3.5),main="Baseline",xlab="")  
 curve(dnorm(x,0,b2\_sig),add=TRUE, col="blue")  
   
 legend(2,2,legend=c("Basline density", "Baseline normal distribution"),col=c("black","blue"),lty=1, cex=0.8)



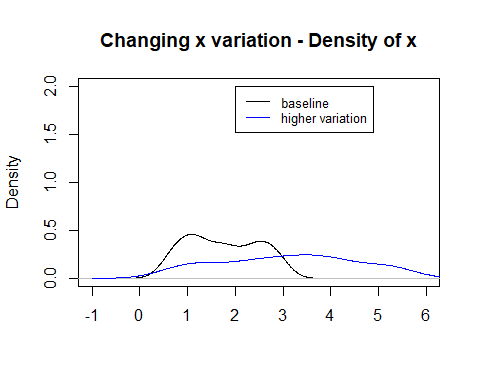
# Changing obs  
 plot(db2,xlim=c(-4,4),main="Changing the number of observations",xlab="")  
 lines(db2\_less\_obs,col="blue")  
 legend(2,.2,legend=c("obs=100", "obs=10"),col=c("black","blue"),lty=1, cex=0.8)



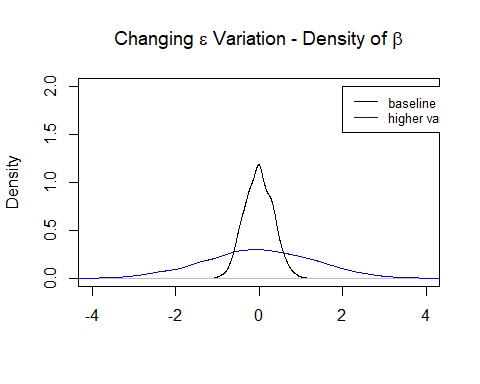
# Changing x variation  
 plot(db2,xlim=c(-1.5,1.5),ylim=c(0,3),main=TeX("Changing x variation - Density of $\\beta$"),xlab="")  
 lines(db2\_more\_var\_x,col="blue")  
 legend(.5,2.5,legend=c("baseline", "higher variation"),col=c("black","blue"),lty=1, cex=0.8)



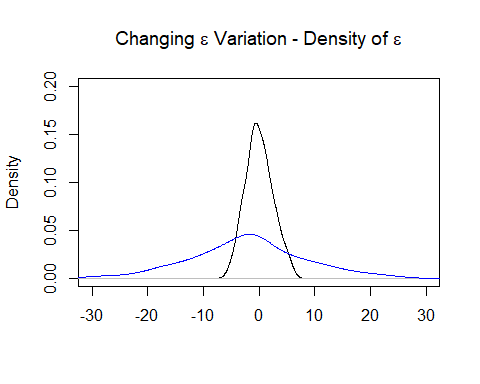
#legend(4,.5,legend=c("baseline", "higher variation"),col=c("black","blue"),lty=1, cex=0.8)  
  
 #summary(x)  
 dx=density(x)  
 dx\_more\_var\_x=density(x\_more\_var\_x)  
 plot(dx,xlim=c(-1,6),ylim=c(0,2),main="Changing x variation - Density of x",xlab="")  
 lines(dx\_more\_var\_x,col="blue")  
 legend(2,2,legend=c("baseline", "higher variation"),col=c("black","blue"),lty=1, cex=0.8)



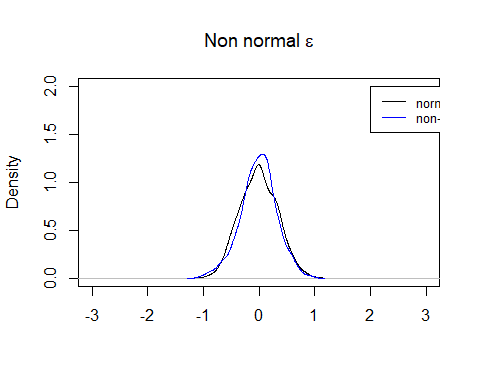
# Changing eps variation  
   
   
 plot(db2,xlim=c(-4,4),ylim=c(0,2),main=TeX("Changing $\\epsilon$ Variation - Density of $\\beta$"),xlab="")  
 lines(db2\_more\_var\_eps,col="blue")  
 legend(2,2,legend=c("baseline", "higher variation"),col=c("black","blue"),lty=1, cex=0.8)



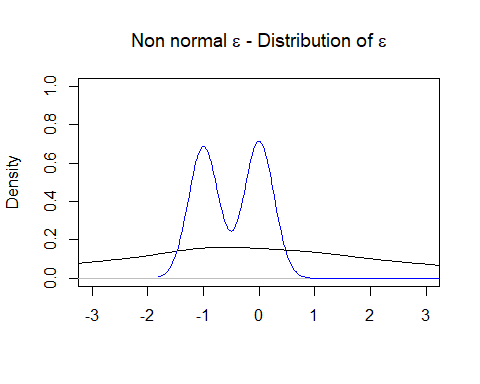
plot(db2\_eps,xlim=c(-30,30),ylim=c(0,.2),  
 main=TeX("Changing $\\epsilon$ Variation - Density of $\\epsilon$"),xlab="")   
 lines(db2\_eps\_more\_var\_eps,col="blue")



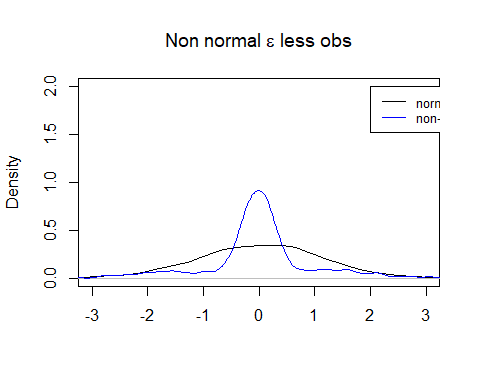
# Non normal eps  
 plot(db2,xlim=c(-3,3),ylim=c(0,2),main=TeX("Non normal $\\epsilon$"),xlab="")  
 lines(db2\_non\_normal\_eps,col="blue")  
 legend(2,2,legend=c("normal", "non-normal"),col=c("black","blue"),lty=1, cex=0.8)



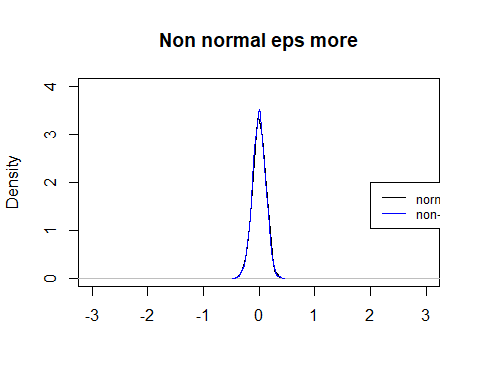
deps=density(eps)  
 deps\_non\_normal\_eps=density(eps\_non\_normal\_eps)  
  
   
 plot(deps,xlim=c(-3,3),ylim=c(0,1),main=TeX("Non normal $\\epsilon$ - Distribution of $\\epsilon"),xlab="")  
 lines(deps\_non\_normal\_eps,col="blue")  
 legend(5,.5,legend=c("normal", "non-normal"),col=c("black","blue"),lty=1, cex=0.8)



# Non normal eps less  
 plot(db2\_less\_obs,xlim=c(-3,3),ylim=c(0,2),main=TeX("Non normal $\\epsilon$ less obs"),xlab="")  
 lines(db2\_non\_normal\_eps\_less,col="blue")  
 legend(2,2,legend=c("normal", "non-normal"),col=c("black","blue"),lty=1, cex=0.8)



# Non normal eps more  
 plot(db2\_more\_obs,xlim=c(-3,3),ylim=c(0,4),main="Non normal eps more",xlab="")  
 lines(db2\_non\_normal\_eps\_more,col="blue")  
 legend(2,2,legend=c("normal", "non-normal"),col=c("black","blue"),lty=1, cex=0.8)



# Hypothesis Testing in the foreigners and crime example

library(haven) # make sure libraries such as this are installed. If not go to Tools -> Install Packages  
 df=ff  
   
 df['crimesPc']=df$crimes11/df$pop11  
 reg1=lm(crimesPc~b\_migr11,df)  
 summary(reg1)

##   
## Call:  
## lm(formula = crimesPc ~ b\_migr11, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.13314 -0.33959 -0.06763 0.22302 2.92572   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.091273 0.045146 24.17 < 2e-16 \*\*\*  
## b\_migr11 0.025164 0.002922 8.61 3.33e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5482 on 321 degrees of freedom  
## Multiple R-squared: 0.1876, Adjusted R-squared: 0.1851   
## F-statistic: 74.14 on 1 and 321 DF, p-value: 3.325e-16

library("car")

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

linearHypothesis(reg1, c( "b\_migr11= 0.04") )

## Linear hypothesis test  
##   
## Hypothesis:  
## b\_migr11 = 0.04  
##   
## Model 1: restricted model  
## Model 2: crimesPc ~ b\_migr11  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 322 104.212   
## 2 321 96.467 1 7.7448 25.771 6.525e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Working out the cut off points

pnorm(-1.644854)

## [1] 0.04999996

pnorm(-1.959964)

## [1] 0.025

qnorm(0.005)  
qnorm(0.025) qnorm(0.05)

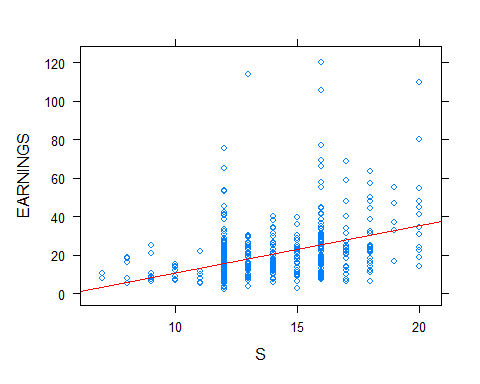
qnorm(0.995)  
qnorm(0.975) qnorm(0.95)

#< Annother example

library(foreign)  
 library(haven)  
 eaef <- read.csv("https://www.dropbox.com/s/31lyn5p5edyoxl5/eaef21.csv?dl=1")  
 head(eaef, n=10)

## X ID FEMALE MALE ETHBLACK ETHHISP ETHWHITE AGE S EDUCPROF EDUCPHD  
## 1 1 5531 0 1 0 0 1 45 12 0 0  
## 2 2 2658 0 1 0 1 0 40 12 0 0  
## 3 3 5365 0 1 0 0 1 38 15 0 0  
## 4 4 4468 0 1 0 0 1 43 13 0 0  
## 5 5 3142 0 1 0 0 1 38 18 0 0  
## 6 6 2170 0 1 1 0 0 39 16 0 0  
## 7 7 2344 0 1 0 0 1 40 14 0 0  
## 8 8 4583 0 1 0 0 1 37 12 0 0  
## 9 9 2517 0 1 0 0 1 39 12 0 0  
## 10 10 3563 0 1 0 0 1 42 12 0 0  
## EDUCMAST EDUCBA EDUCAA EDUCHSD EDUCDO SINGLE MARRIED DIVORCED FAITHN FAITHP  
## 1 0 0 0 1 0 0 1 0 0 1  
## 2 0 0 0 0 1 0 0 1 1 0  
## 3 0 0 1 0 0 0 0 1 0 0  
## 4 0 0 0 1 0 0 1 0 0 1  
## 5 1 0 0 0 0 0 1 0 0 1  
## 6 0 0 0 0 0 0 1 0 0 1  
## 7 0 0 1 0 0 1 0 0 0 1  
## 8 0 0 0 1 0 1 0 0 0 1  
## 9 0 0 0 1 0 0 1 0 0 1  
## 10 0 0 0 1 0 0 1 0 0 1  
## FAITHC FAITHJ FAITHO ASVAB01 ASVAB02 ASVAB03 ASVAB04 ASVAB05 ASVAB06  
## 1 0 0 0 58 64 52 56 54 56  
## 2 0 0 0 32 39 29 29 27 22  
## 3 0 0 1 42 40 37 38 42 45  
## 4 0 0 0 66 55 59 53 51 52  
## 5 0 0 0 64 65 61 62 47 34  
## 6 0 0 0 62 53 51 59 58 45  
## 7 0 0 0 50 32 44 38 33 29  
## 8 0 0 0 56 62 52 44 41 47  
## 9 0 0 0 60 59 55 62 58 46  
## 10 0 0 0 46 57 28 32 47 46  
## ASVABC HEIGHT WEIGHT85 WEIGHT02 SM SF SIBLINGS LIBRARY POV78 EXP  
## 1 60.89985 67 160 200 8 8 11 0 0 22.384615  
## 2 33.63790 67 185 205 5 5 3 0 1 8.903846  
## 3 38.81767 69 135 185 11 12 3 1 0 13.250000  
## 4 57.08318 72 200 250 12 16 2 1 0 18.250000  
## 5 65.53439 76 185 220 16 20 1 1 0 13.769231  
## 6 55.44746 69 180 215 12 12 2 0 0 11.692307  
## 7 36.36409 65 150 145 12 12 1 1 0 19.480770  
## 8 56.53794 69 125 140 8 13 10 1 0 15.230769  
## 9 60.62724 72 195 240 12 12 3 1 NA 14.365385  
## 10 43.99744 70 180 210 10 8 9 1 0 22.038462  
## EARNINGS HOURS TENURE COLLBARG CATGOV CATPRI CATSE URBAN REGNE REGNC REGW  
## 1 53.41 45 2.7500000 0 0 1 0 0 0 0 0  
## 2 8.00 40 2.3846154 0 0 1 0 0 0 0 0  
## 3 24.00 40 5.7500000 1 0 1 0 1 0 1 0  
## 4 29.50 40 6.1346154 1 0 1 0 1 0 0 1  
## 5 32.05 54 0.8269231 0 1 0 0 1 1 0 0  
## 6 14.73 40 4.2884617 0 0 1 0 1 0 0 0  
## 7 13.00 48 0.7692308 0 0 1 0 1 0 0 0  
## 8 11.25 40 5.1538463 0 0 1 0 1 0 0 1  
## 9 18.01 40 1.5769231 1 1 0 0 1 0 1 0  
## 10 11.78 40 0.2307692 1 0 1 0 1 0 1 0  
## REGS  
## 1 1  
## 2 1  
## 3 0  
## 4 0  
## 5 0  
## 6 1  
## 7 1  
## 8 0  
## 9 0  
## 10 0

library("lattice")  
 xyplot(EARNINGS ~ S, data = eaef, type = c("p","r"), col.line = "red")



mod\_earn <- lm(EARNINGS ~ S , data = eaef)  
 summary(mod\_earn)

##   
## Call:  
## lm(formula = EARNINGS ~ S, data = eaef)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -24.112 -7.142 -2.476 3.625 95.974   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -13.9335 3.2199 -4.327 1.8e-05 \*\*\*  
## S 2.4553 0.2319 10.590 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.13 on 538 degrees of freedom  
## Multiple R-squared: 0.1725, Adjusted R-squared: 0.171   
## F-statistic: 112.1 on 1 and 538 DF, p-value: < 2.2e-16

#>

# tea time

qt(0.025, 10 )

## [1] -2.228139

# Effect of Experience on Earnings

library(foreign)  
 eaef <- read.csv("https://www.dropbox.com/s/31lyn5p5edyoxl5/eaef21.csv?dl=1")  
  
 mod\_earn\_exp <- lm(EARNINGS ~ EXP , data = eaef)   
 summary(mod\_earn\_exp)

##   
## Call:  
## lm(formula = EARNINGS ~ EXP, data = eaef)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -17.140 -8.876 -3.723 3.869 99.986   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 15.5553 2.4425 6.369 4.09e-10 \*\*\*  
## EXP 0.2415 0.1398 1.727 0.0847 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.39 on 538 degrees of freedom  
## Multiple R-squared: 0.005515, Adjusted R-squared: 0.003666   
## F-statistic: 2.983 on 1 and 538 DF, p-value: 0.08469

# More general hypothesis tests

mod\_exp\_s <- lm(EXP ~ S , data = eaef)  
 summary(mod\_exp\_s)

##   
## Call:  
## lm(formula = EXP ~ S, data = eaef)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -17.0512 -2.3320 0.8564 3.1391 6.3756   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 22.3165 1.0624 21.006 < 2e-16 \*\*\*  
## S -0.3961 0.0765 -5.178 3.17e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.331 on 538 degrees of freedom  
## Multiple R-squared: 0.04748, Adjusted R-squared: 0.04571   
## F-statistic: 26.82 on 1 and 538 DF, p-value: 3.169e-07

library(data.table)

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

library("car")  
 linearHypothesis(mod\_exp\_s, c( "S = -1") )

## Linear hypothesis test  
##   
## Hypothesis:  
## S = - 1  
##   
## Model 1: restricted model  
## Model 2: EXP ~ S  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 539 11260   
## 2 538 10091 1 1168.7 62.307 1.658e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1