**EXAMPLE** A researcher wants to test if the population standard deviation is greater than 100. He takes a random sample of 25 values from the population:

0.0	0.0	0.0000068	0.0000135	0.0000176	0.0000209	0.0003327	0.007412	0.0147
0.03630	0.08643	0.08937	0.4875	0.8276	1.406	2.685	3.294	
3.391	4.989	9.712	12.01	37.29	169.7	247.1	598.1	

The data produces a standard deviation of 129.36 which would seem to indicate that σ is greater than 100, that is, $H_1 : \sigma > 100$. Based on the normal reference plot and the Shapiro-Wilk test there is strong evidence that the data is not normally distributed.

Normal Reference Distribution Plot
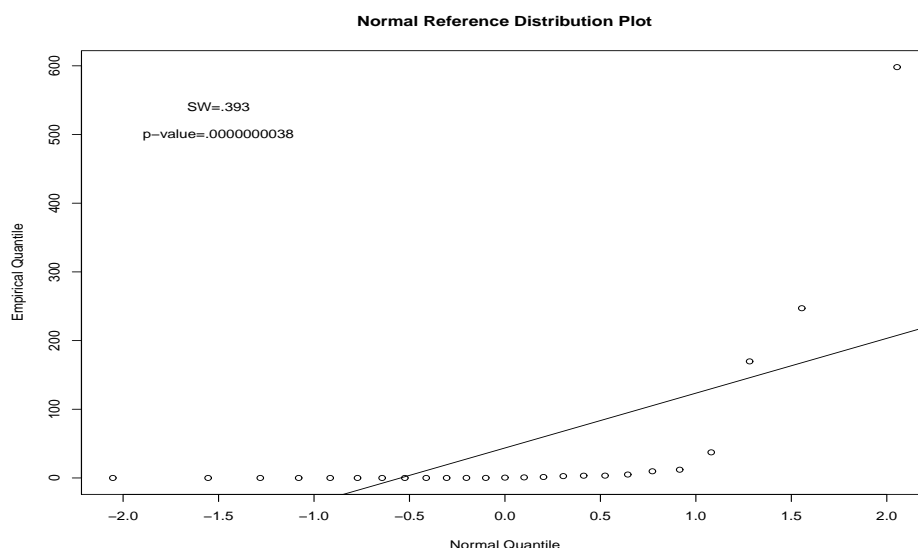
The researcher ignores the non-normality and calculates a lower bound for σ and a p-value for the test using the normal based procedure:

$$\frac{(n-1)S^2}{\sigma_o^2} = \frac{(25-1)(129.36)^2}{(100)^2} = 40.16 \ \Rightarrow \ p - value = 1 - pchisq(40.16, 24) = .021$$

The normal based lower 95% confidence bound for σ is $\sqrt{\frac{(n-1)S^2}{\chi^2(.05,24)}} = 105.02$ which is greater than 100 and hence would support the conclusion that σ is greater than 100.

The company's statistician informs the researcher that it would be a good idea to use a bootstrap procedure because the test statistics is highly sensitive to a lack of normality. The following R code produces a bootstrap a p-value for testing $H_o : \sigma \leq 100$ versus $H_1 : \sigma > 100$. Recall that the p-value is the probability of obtaining a value of the test statistic larger than the value of the test statistic from the data set under the assumption that H_o is true, that is, $\sigma = 100$. Thus, it is necessary to transform the data into a data set having a standard deviation of 100 prior to conducting the resampling. Transform the data, X's, by $Y = 100 * X/sd(X)$ which yields a data set having $sd(Y) = 100$

The following R code (bootTestofSigma.R) yields the bootstrap p-value and a bootstrap lower 95% confidence bound on σ:

```
x = c(0.0,0.0,0.0000068,0.0000135,0.0000176,0.0000209,0.0003327,0.007412,
      0.0147,0.03630,0.08643,0.08937,0.4875,0.8276,1.406,2.685,3.294,3.391,4.989,
      9.712,12.01,37.29,169.7,247.1,598.1)


n = length(x)
x = sort(x)
i = seq(1:n)
u = (i-.5)/n
z = qnorm(u)
plot(z,x,datax=F,plot=T,xlab="Normal Quantile",ylab="Empirical Quantile",
      lab=c(7,8,7),
      main="Normal Reference Distribution Plot",
      cex=.95)
abline(lm(x~z))
shapiro.test(x)
text(-1.5,540,"SW=.393")
text(-1.5,500,"p-value=.0000000038")
m = mean(x)
SD = sd(x)
sigma0 = 100
TSD = (n-1)*SD^2/sigma0^2
B  =  9999
PV =  numeric(B)
PV =  rep(0,B)
TS =  numeric(B)
TS =  rep(0,B)
S =  numeric(B)
S =  rep(0,B)
Sy =  numeric(B)
Sy =  rep(0,B)


{
for (i in 1:B)
S[i]  =  sd(sample(x,replace=T))
}


PV = (n-1)*S^2/SD^2
PV =  sort(PV)


#standardize the data so that the null hypothesis is true


y = x*sigma0/sd(x)


{
for (i in 1:B)
Sy[i]  =  sd(sample(y,n,replace=T))
}
```

```
TS = (n-1)*Sy^2/sigma0^2

SIMpvalue = sum(TS>TSD)/B

NORMpvalue = 1-pchisq(TSD,24)

LPV  =  PV[250]
UPV  =  PV[9750]

LPV2 = PV[9500]

SIMci = c(sqrt((n-1)*SD^2/UPV), sqrt((n-1)*SD^2/LPV))

SIMLcb = sqrt((n-1)*SD^2/LPV2)

NORMLcb = sqrt((n-1)*SD^2/qchisq(.95,n-1))



SD
# 129.3631

SIMpvalue
# .2022

NORMpvalue
# .0202

SIMLcb
# 84.15

NORMLcb
105.02
```

Note that the bootstrap lower 95% confidence bound is 84.15 which is less than 100 and the bootstrap p-value is .2022 both of which imply that there is not significant evidence that σ is greater than 100.