عنوان پروژه: نمودار های احتمال و بررسی توزیع داده‌ها

پدید آورنده: محراب عتیقی

نمودار های احتمال و بررسی توزیع داده‌ها

یکی از کاربردهای آنها، بررسی کفایت یک مدل آماری است. بدین معنی که بر اساس داده هایی که از جامعه به دست می‌آوریم، نمودار احتمال را می‌توان برای بررسی اینکه آیا داده ها از یک مدل خاص آماری استخراج شده اند به کار برد.

کاربرد دیگر این گونه نمودارها، پس از تعیین مدل، برآورد پارامترها و چندک های توزیع آماری است.

از نقاط قوت نمودارهای احتمال می توان به موارد زیر اشاره کرد:

- ساده و سریع بودن در استفاده، در مقابل روش‌های دیگر استنباط آماری ممکن است دارای محاسبات پیچیده و استفاده از آنها مستلزم روش آماری نسبتاً بیشتری باشد.
- نمودارهای احتمال روشی ساده برای بررسی کفایت یک توزیع، برآورد پارامترهای توزیع، چندک ها و دیگر مشخصه های جمعه مورد بررسی فراهم می کند.
- نمودارهای احتمال برای هر دو حالت داده های **کامل** و داده های **سانسور** شده به کار می‌روند.
- نمودارهای احتمال می توانند برای بررسی فرضیات روش‌های استنباط پارامتری قبل از به کار بردن چنین روش‌هایی مفید واقع شوند.

اما در کنار چنین نقاط قوتی دارای ضعف‌هایی نیز می باشند که از جمله می‌توان به ذهنی بودن آنها اشاره کرد. بدین معنی که ممکن است دو فرد برداشت متفاوتی از یک نمودار احتمال داشته باشند. همچنین از برآوردهای به دست آمده از روش نمودار احتمال نمی توان فاصله اطمینان به دست آورد.

برآورد تابع توزیع

اگر F یک تابع توزیع احتمال باشد، آنگاه برآورد ناپارامتری آن بر اساس یک نمونه تصادفی به صورت زیر تعریف شد که به آن برآوردگر تجربی تابع توزیع می‌گویند.

$$\hat{F}\,(t) = \frac{\text{تعداد مشاهداتی که در نمونه کمتر یا مساوی } t \text{ هستند}}{n}$$

اگر t_1وt_2و...وt_n مقادیر مرتب شده نمونه باشند،آنگاه باتوجه به این تعریف داریم:

$$F_n(t_i) = \frac{i}{n}$$

یعنی احتمال اینکه طول عمر کمتر یا مساوی t_i باشد با کمیت $\frac{i}{n}$ برآورد می‌شود. در نظریه احتمال ثابت می‌شود که هنگامی که n به سمت بی‌نهایت میل می‌کند آنگاه $\hat{F}_n(t)$ به سمت F(t) میل می‌کند بنابراین هرچه قدر حجم نمونه بیشتر باشد آنگاه $\hat{F}_n(t)$ به F(t) نزدیک خواهد شد. پس به طور تقریب می‌توان برای حجم نمونه‌هایی به اندازه کافی بزرگ نوشت،

$$F(t_i) \cong \hat{F}_n(t_i) = \frac{i}{n} \quad , i = 1,2, \dots, n$$

از این رابطه می‌توان تقریب زیر را نوشت،

$$F^{-1}\left(\frac{i}{n}\right) \cong t_i \, , i = 1,2, \dots, n$$

اگر نمونه تصادفی از توزیع F استخراج شده باشدو t_1وt_2و...وt_n را مقادیر مرتب شده نمونه در نظر بگیریم.
آنگاه انتظار داریم که نقاط $(F^{-1}\left(\frac{i}{n}\right),t_i)$ ، $i = 1,2,\ldots,n$ حول یک خط راست قرار بگیرند که از مبدا می گذرد و ضریب زاویه آن ۱ است. بنابراین، به راحتی به شیوه‌ای دست یافته ایم که با استفاده از آن می‌توان کفایت یک مدل را بررسی کرد. یک نمونه تصادفی از متغیر طول عمر مورد نظر اختیار کرده و مقادیر نمونه را مرتب می‌کنیم. فرض می‌کنیمF یک توزیع مفروض باشد که علاقه‌مندیم آزمون کنیم که آیا داده ها از آن استخراج شده اند.

برای انجام این کار نقاط $\left(F^{-1}\left(\frac{i}{n}\right), t_i\right)$ را روی صفحه مختصات رسم می‌کنیم. اگر نقاط رسم شده حول یک خط راست باشند که ضریب زاویه آن ۱ و از مبدا عبور می‌کند آنگاه می پذیریم که داده از توزیع F آمده اند.

نمودار احتمال p-p

نمودار P- P

نمودارP- P نیز همانند نمودارQ-Q برای بررسی این است که داده از یک توزیع خاص آمده باشند. در این نمودار فراوانی تجمعی نسبی مشاهده ها و احتمال های تجمعی توزیع مد نظر به صورت زوج هایی محاسبه شده و در یک دستگاه مختصات رسم می‌شوند.

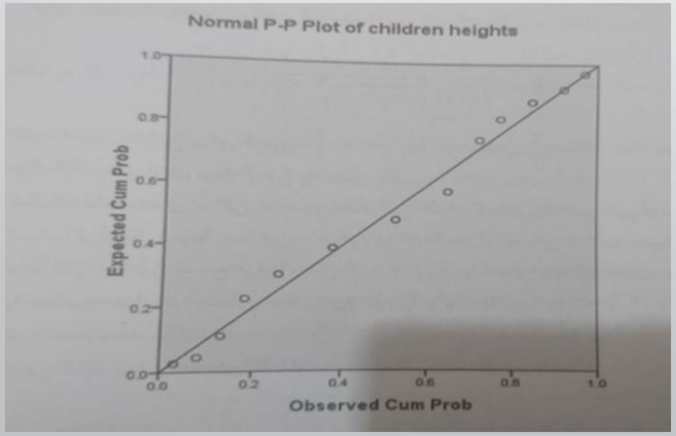
در این دستگاه فراوانی تجمعی نسبی مشاهده ها در روی محور افقی به عنوان مقادیر مشاهده شده و احتمال های تجمعی توزیع مد نظر بر روی محور عمودی به عنوان مقادیر مورد انتظار درج می‌گردد.

نمودار احتمال p-p

در صورتی که داده های مشاهده شده از توزیع نرمال باشند، نقاط حاصل از زوج های فراوانی تجمعی نسبی مشاده ها و احتمال تجمعی توزیع نرمال بایستی روی یک خط مستقیم که از نقاط (0,0) و (1,1) می‌گذرند قرار گیرند. اگر این نقاط نزدیک این خط مستقیم باشند، داده ها از توزیع نرمال هستند و در غیر این صورت داده ها از توزیع نرمال نیستند.

بعنوان مثال در شکل بعدی مشاهده می‌شود که نقاط نزدیک به خط مستقیم هستند، پس داده ها از توزیع نرمال پیروی می‌کنند.

در این بخش برای بررسی نرمال بودن داده ها از نمودار هایP- P و Q-Q استفاده کردیم.



نتیجه گیری از روی این نمودار ها تنها بر اساس مشاهده نمودار و بیان اینکه نقاط به خط مستقیم نزدیک هستندیا نه صورت می‌گیرد. اما این نمودار ها معیاری را برای نزدیک یا دور بودن نقاط از خط ارائه نمی‌دهند. برای بررسی دقیق تر نرمال بودن داده ها می‌توان از آزمون های برازندگی توزیع که بدین منظور تهیه شده اند استفاده کرد . دو نوع از این آزمون ها، آزمون خی دو برای برازندگی توزیع و آزمون کولموگوروف _ اسمیرنوف است.

نمودار احتمال توزیع نمایی

فرض کنید توزیع مورد بررسی نمایی باشد با تابع توزیع احتمال

$$F(t) = 1 - e^{-\lambda(t-\theta)} \ , \qquad t > \theta \, , \theta \, , \lambda > 0$$

مراحل تشکیل نمودار احتمال و برآورد پارامترهای توزیع احتمال بصورت زیر است:

۱) نقاط t_i را در مقابل $E_i = -\ln(1 - \frac{i}{n+1})$ و $i = 1,2, \dots ,n$ رسم می‌کنیم. اگر برازش مناسب باشد.

۲) λ, θ بصورت زیر برآورد می‌شوند:

$$\hat{\lambda} = \frac{1}{\text{ضریب زاویه}} \quad \text{و} \quad \hat{\theta} = \text{عرض از مبدا}$$

نمودار احتمال توزیع نمایی

مثال: فرض کنید ۲۰ مولد برق را در یک آزمون شتابنده وارد کرده ایم. آزمون پس از شکست تمام مولدها خاتمه پیدا کرده و مشاهدات زیر بر حسب ساعت بدست آمده اند.

۷۱۱/۵و۱۰۵۱و۶۳۰۳/۹و۱۸۸۳/۶و۶۰۵۴/۳و۶۸۵۳/۷و۷۲۰۱/۹و۲۷۹/۸و۲۳۱۱/۱و۷/۵و۵۲۹۶/۶و۸۴۸/۲و

۹۰۶۸/۵و۱۰۶۰۹/۷و۵۹۲/۱و۱۶۵۷/۲و۵۶۳۷/۹و۲۹۵۱/۹و۱۴۲۵/۵و۱۲۱/۵

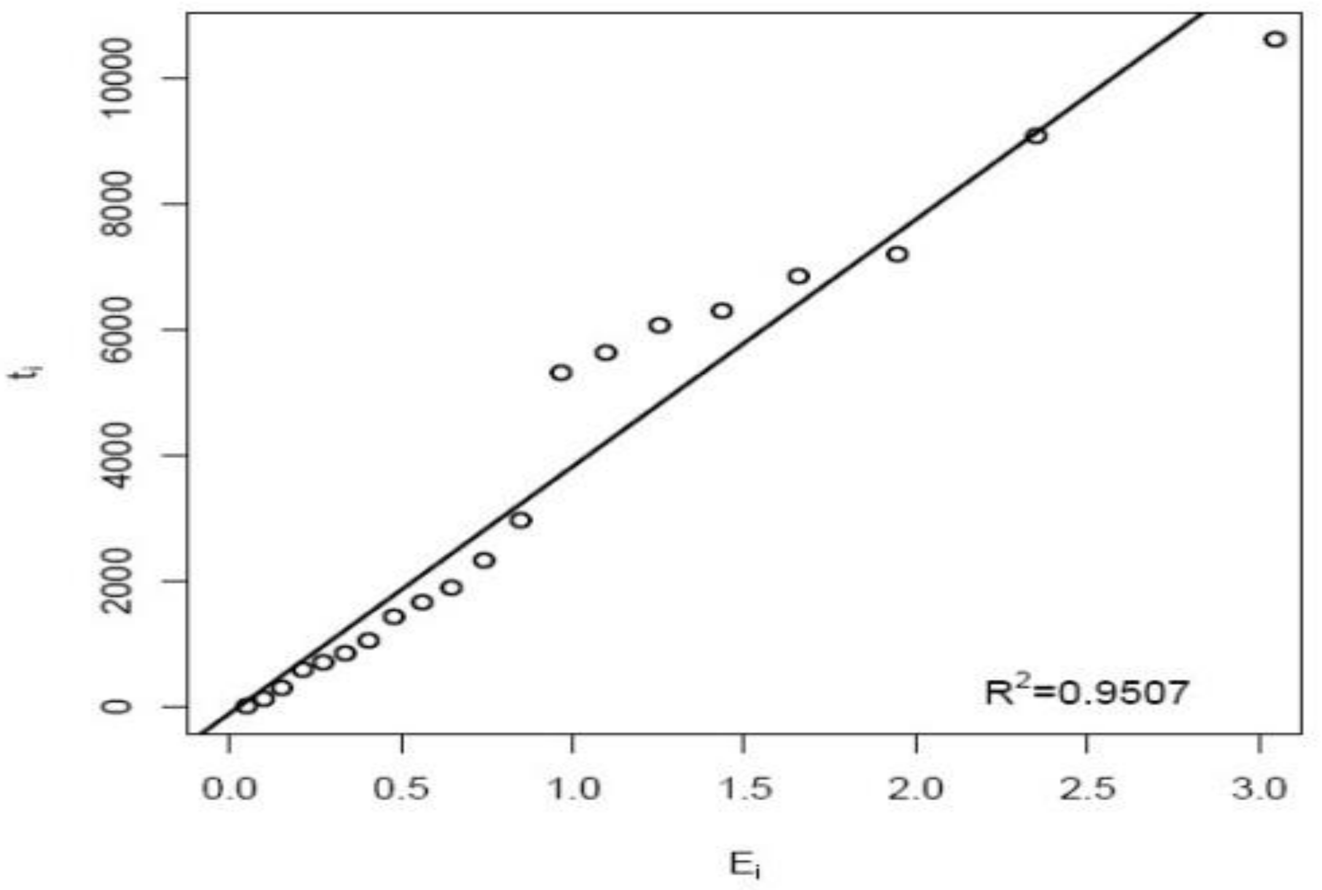
بررسی کنید که آیا توزیع نمایی برازش مناسبی برای داده ها است؟ در صورت مثبت بودن پارامترهای توزیع را برآورد کنید.

نمودار احتمال توزیع نمایی

حل. درشکل ۸ . ۱نمودار احتمال این مقادیر برای توزیع نمایی نمایش داده شده است یعنی مقادیر مرتب شده t_i را در مقابل E_i و $i = 1,2, \ldots ,n$ رسم می‌کنیم.از نمودار بنظر می‌رسد برازش توزیع نمایی مناسب است.خط کمترین مربعات که از داده ها می‌گذرد دارای عرض از مبدا ۱۰۳/۱ − ضریب زاویه ۳۹۳۰/۴ است بنابراین برآورد λ, θ به ترتیب برابر با ۰ و ۰/۰۰۰۲۵۴ است.

(اسلاید بعد شکل)

نمودار احتمال توزیع نمایی



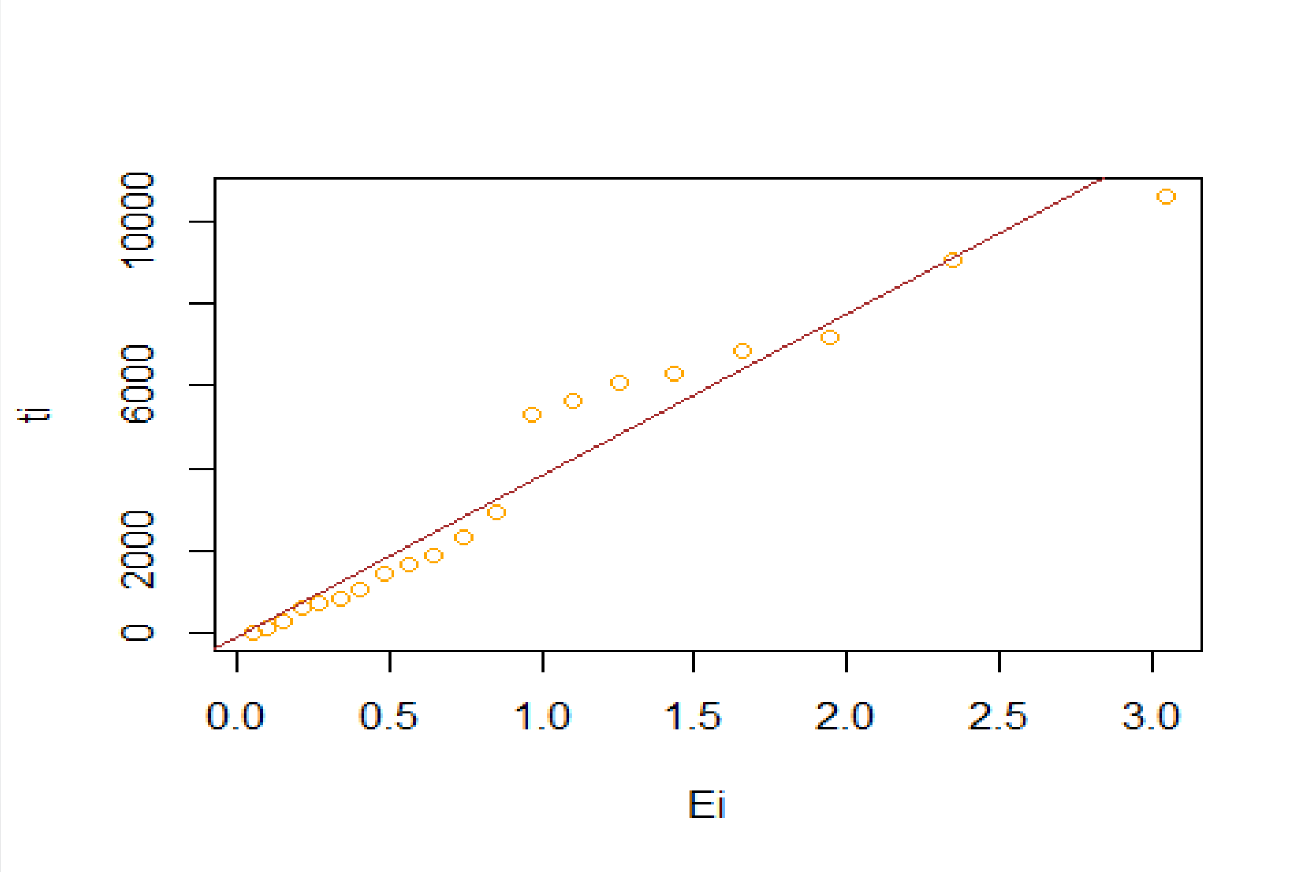
شکل ۱.۸ نمودار احتمال توزیع نمایی برای داده‌های مولد برق

نمودار احتمال توزیع نمایی

خروجی نرم افزار R

```
library(CircStats)

## Warning: package 'CircStats' was built under R version 4.0.5

## Loading required package: MASS

## Loading required package: boot

#example 1.8:
t<-c(711.5,1051,6303.9,1883.6,6054.3,6853.7,7201.9,279.8,2311.1,7.5,5296.6
,848.2,9068.5,10609.7,592.1,1657.2,5637.9,2951.2,1425.5,121.5)
#q-qplot first method:
sample.quantiles.exp<-sort(t)
i<-1:length(t)
E<--log(1-(i/(length(t)+1)))
theoretical.quantiles.exp=E
fit<-lm(sort(t)~E)
plot(E,sort(t),col="Orange",xlab = "Ei",ylab="ti")
abline(fit,col="Brown")
```

نمودار احتمال توزیع نمایی

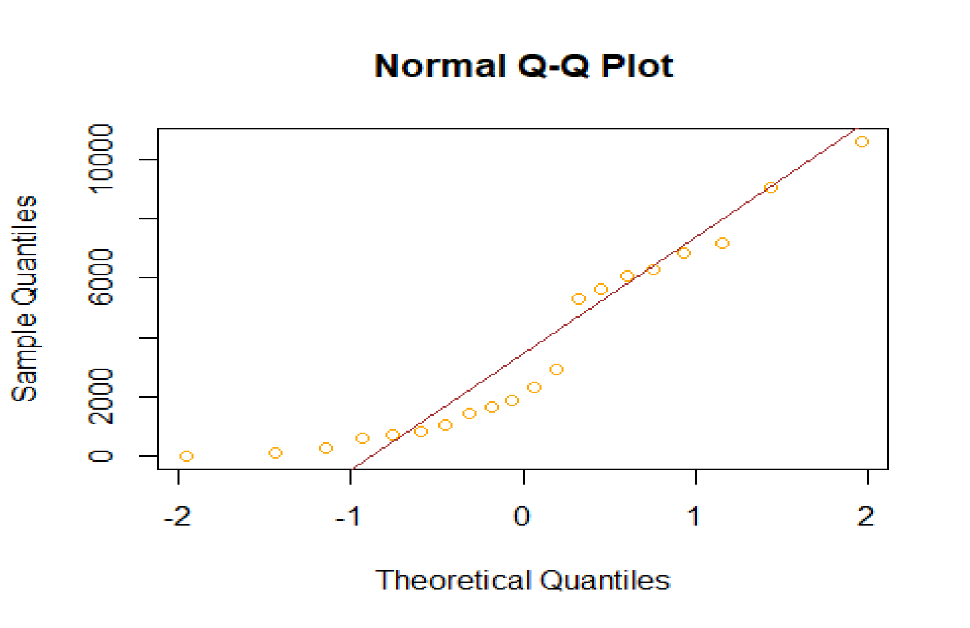



```
fit$coefficients

## (Intercept)           E
##   -103.0802   3930.4210
```

نمودار احتمال توزیع نمایی

```
lamba.hat=1/fit$coefficients[2];theta.hat=fit$coefficients[1]
print(paste("theta.hat is equal to ",theta.hat,
            "the lambda.hat is equal to",lamba.hat))

## [1] "theta.hat is equal to  -103.080151605163 the lambda.hat is equal t
o 0.000254425669078671"

# q-qplot second method:
qqnorm(t,col="Orange")
qqline(t,col="brown")
```

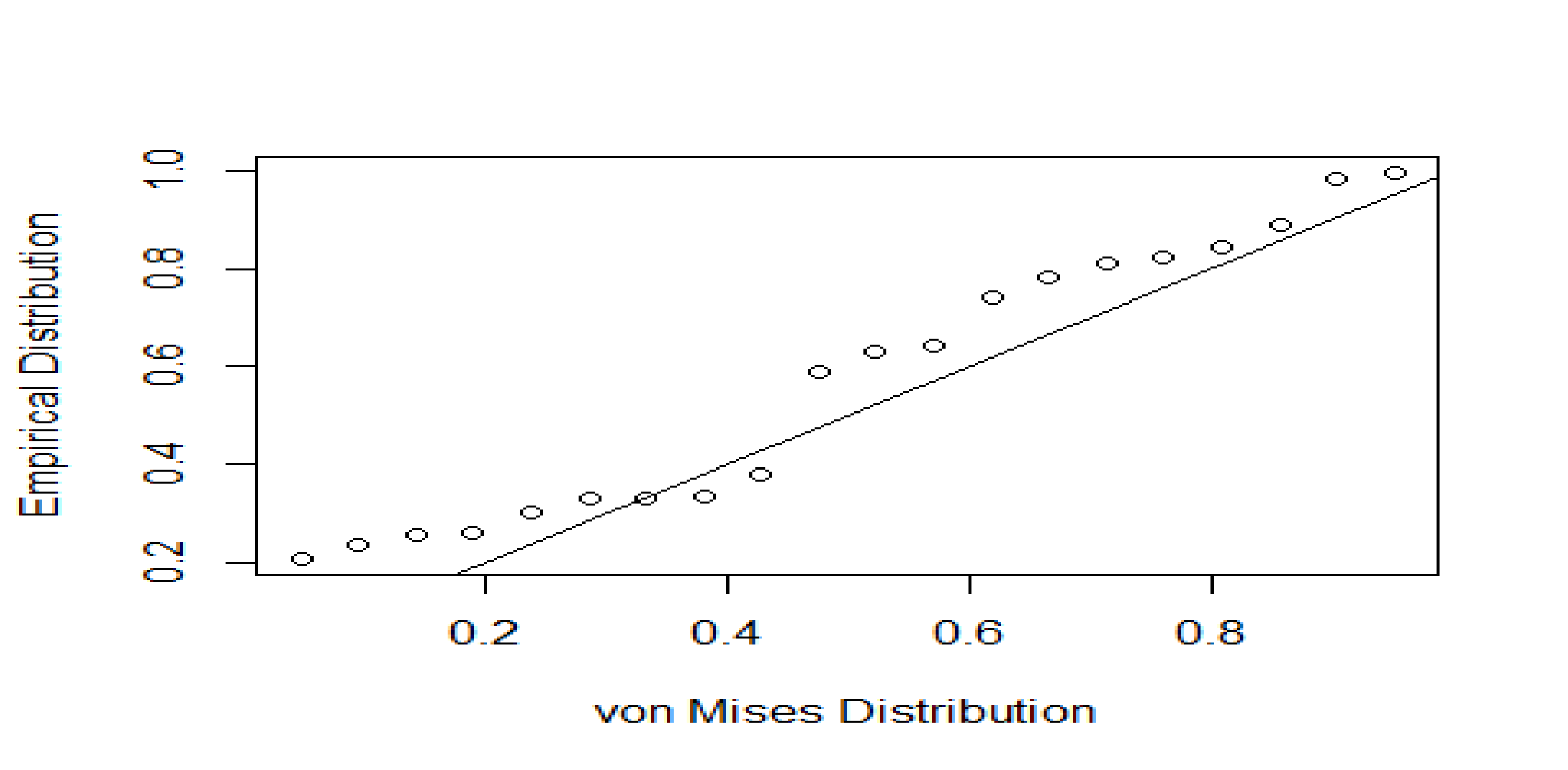
نمودار Q-Q



نمودار احتمال توزیع نمایی

نمودار P-P

```
# p-pplot :
pp.plot(t,ref.line = TRUE)
```




```
##          mu     kappa
## 1 1.688506 0.186204
```

نمودار احتمال توزیع نمایی

خروجی SPSS

Case Processing Summary

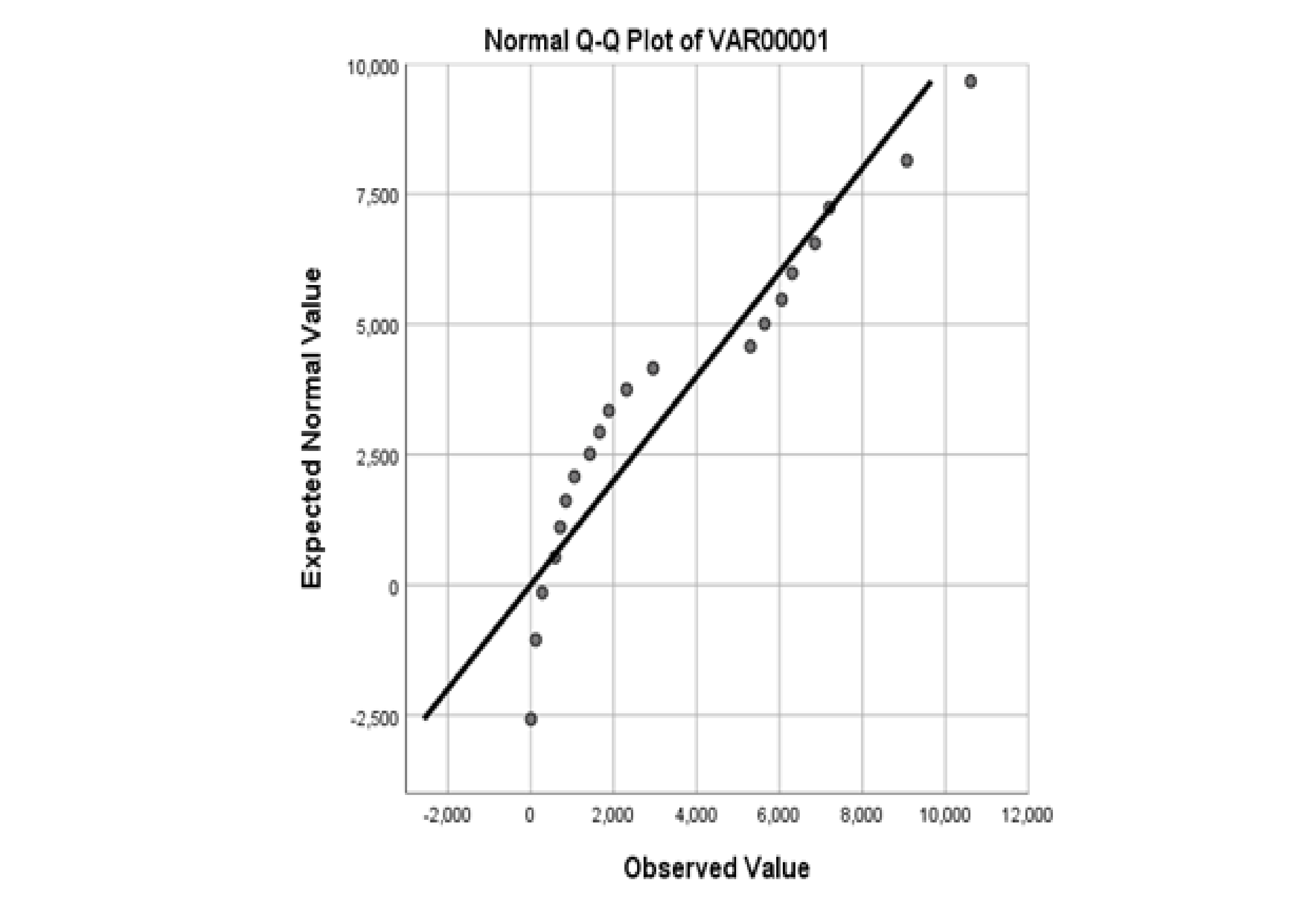
		VAR00001
Series or Sequence Length		20
Number of Missing Values in the Plot	User-Missing	0
	System-Missing	0

Estimated Distribution Parameters

		VAR00001
Normal Distribution	Location	3543.3350
	Scale	3276.10301

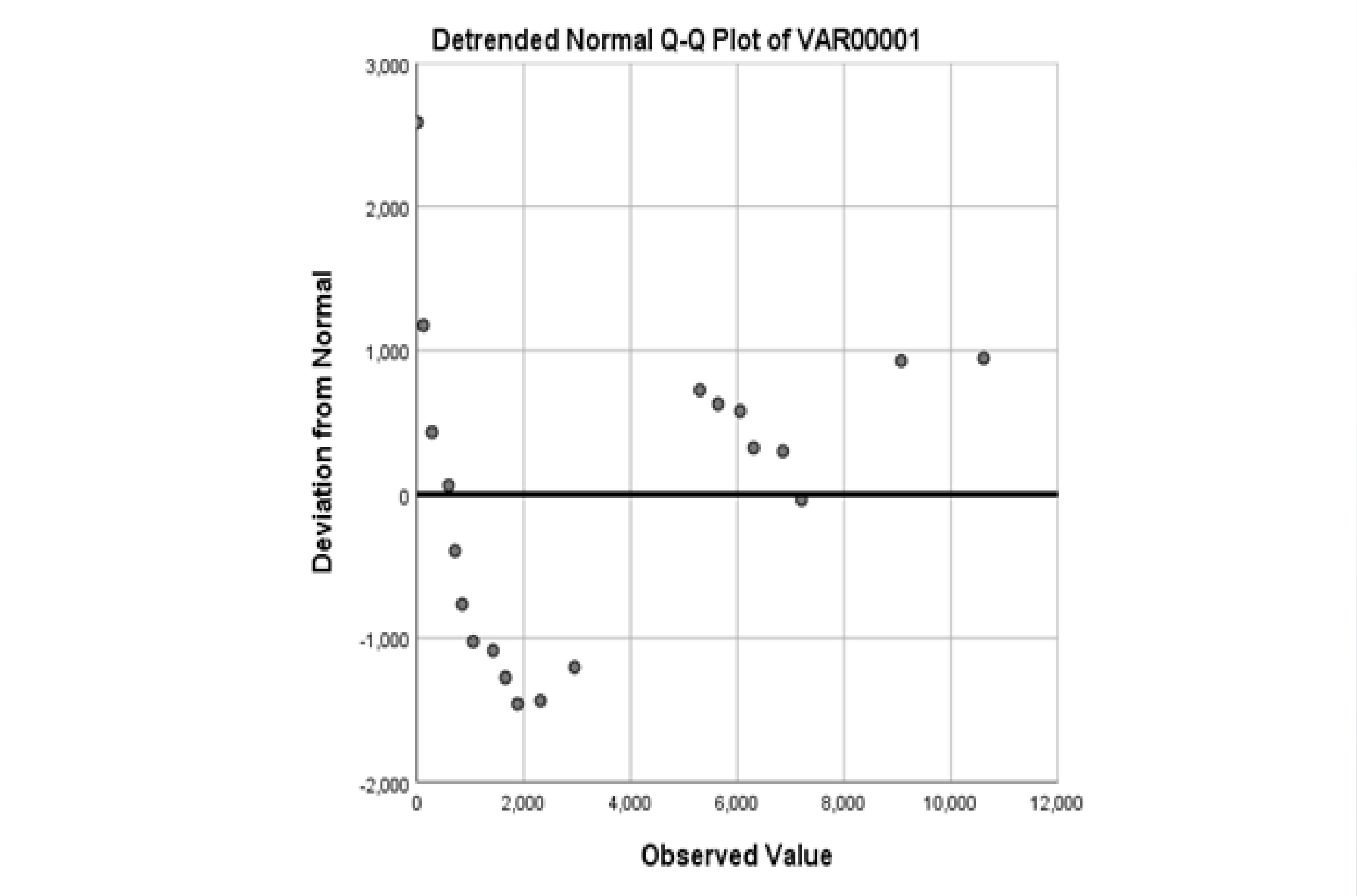
نمودار احتمال توزیع نمایی

نمودار Q-Q



نمودار احتمال توزیع نمایی

نمودار Q-Q



نمودار احتمال توزیع نمایی

خروجی SPSS

Case Processing Summary

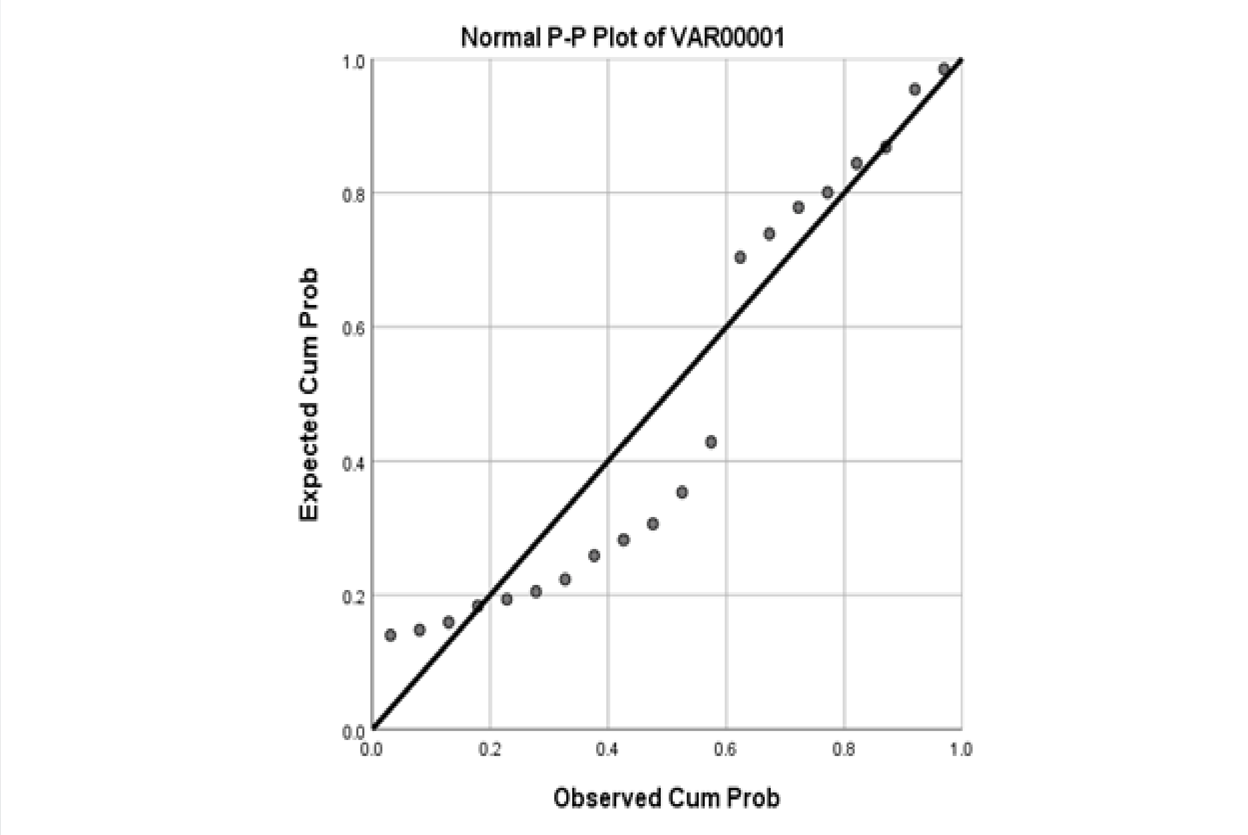
		VAR00001
Series or Sequence Length		20
Number of Missing Values in the Plot	User-Missing	0
	System-Missing	0

Estimated Distribution Parameters

		VAR00001
Normal Distribution	Location	3543.3350
	Scale	3276.10301

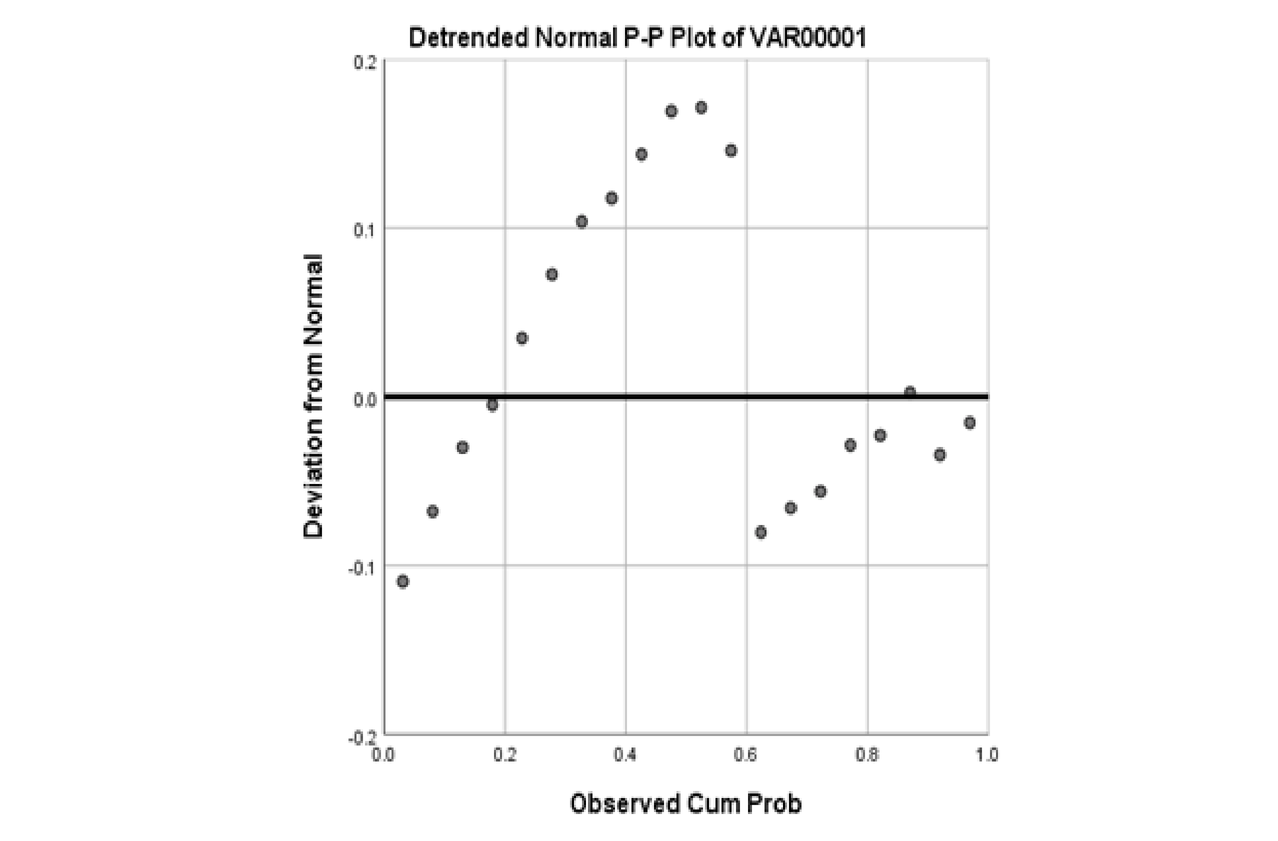
نمودار احتمال توزیع نمایی

نمودار P-P



نمودار احتمال توزیع نمایی

نمودار P-P



نمودار احتمال توزیع های وایبل و مقدار غایی

فرض کنید توزیع طول عمر مورد بررسی وایبل با تابع توزیع زیر باشد:

$$F(t) = 1 - e^{-(\lambda t)^{\beta}} \quad , \qquad t > 0, \lambda, \beta > 0$$

در اینصورت

۱) $\ln t_i$ را در مقابل $W_i = \ln(-\ln(1 - \frac{i}{n+1}))$ رسم می‌کنیم.

۲) برآورد پارامترهای λ,β با استفاده از معادله خط عبارتند از :

$$\lambda = e^{\,-\left(\text{عرض از مبدا}\right)} \quad , \qquad \hat{\beta} = \frac{1}{\text{ضریب زاویه}}$$

نمودار احتمال توزیع های وایبل و مقدار غایی

مثال. فرض کنید مشاهدات زیر، زمان شکست (برحسب سال) یک نوع قطعه هواپیما باشد که در یک آزمون طول عمر به دست آمده اند.

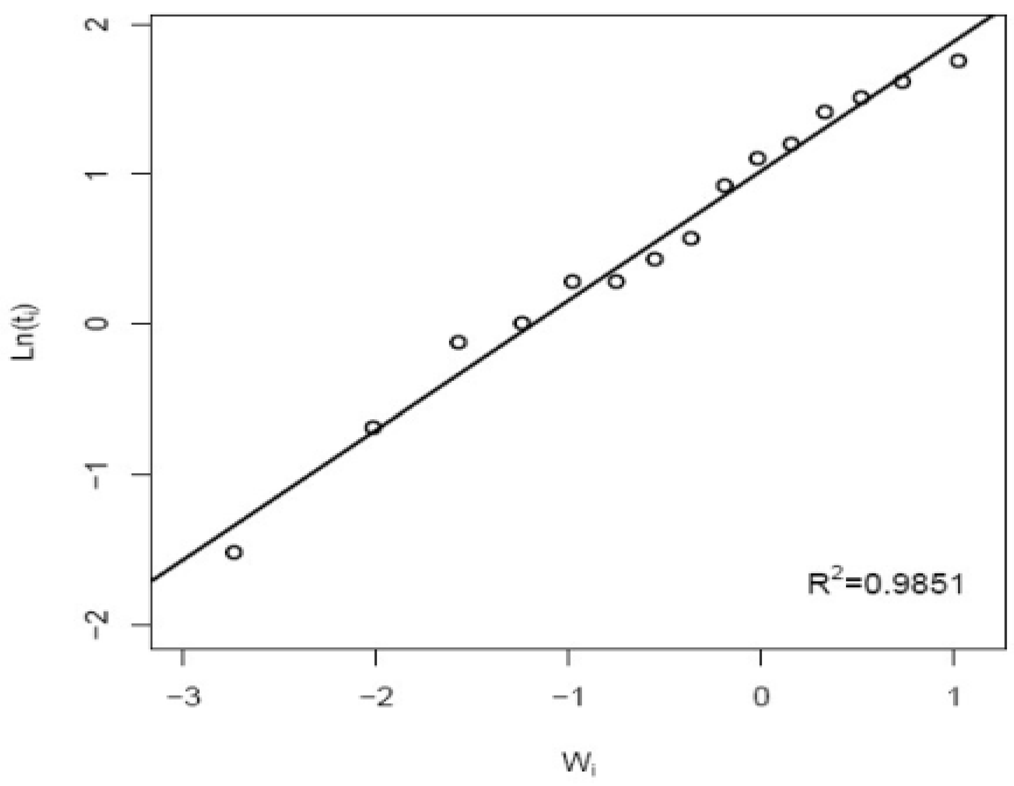
۵/۷۷و۵/۰۳و۴/۵و۴/۱و۳/۳و۳/۰و۲/۵و۱/۷۶و۱/۵۴و۱/۳۳و۱/۳۲و۱و۰/۸۸و۰/۵و۰/۲۲

آیا می‌توان پذیرفت که توزیع طول عمر قطعات وایبل است؟ در صورت مثبت بودن پارامترهای توزیع را بدست آورید.

نمودار احتمال توزیع های وایبل و مقدار غایی

حل. نمودار احتمال داده ها برای توزیع وایبل در شکل ۲.۸ ارائه شده است.باتوجه به اینکه نقاط حول خط برازش شده است و مقدار مربع R مقدار ۰/۹۸۶ است که به مقدار یک نزدیک است بنظر می‌رسد مدل وایبل برای داده ها مدلی مناسب است. خط برازش داده شده دارای عرض از مبدا ۱/۰۲۳ و ضریب زاویه ۰/۸۶۲ است. بنابراین برآورد پارامترهای λ و β به ترتیب ۰/۳۶و ۱/۱۶ است.

نمودار احتمال توزیع های وایبل و مقدار غایی



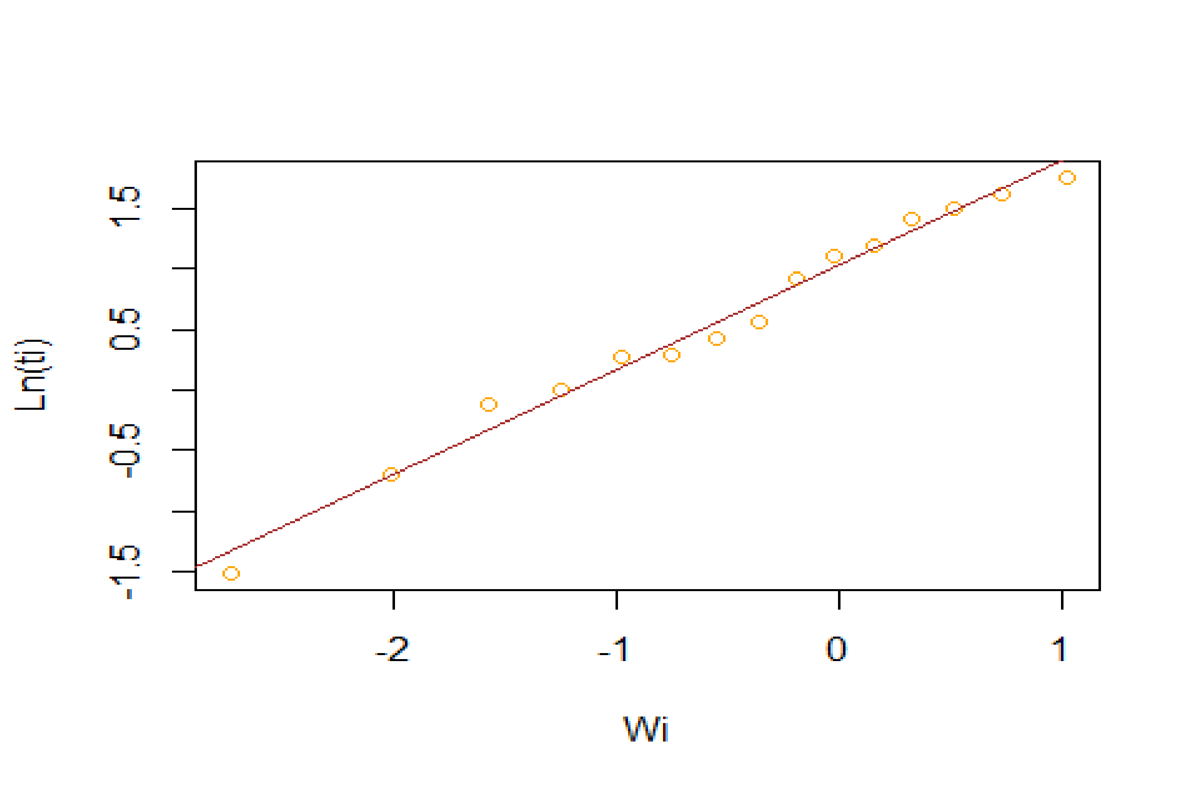
شکل ۲.۸ نمودار احتمال توزیع وایبل برای داده‌های قطعه هواپیما

نمودار احتمال توزیع های وایبل و مقدار غایی

خروجی نرم افزار R

```
#example 2.8:
library("fitdistrplus","qualityTools")

## Warning: package 'fitdistrplus' was built under R version 4.0.4

## Loading required package: survival

##
## Attaching package: 'survival'

## The following object is masked from 'package:boot':
##
##     aml

x<-(c(5.77,5.03,4.5,4.1,3.3,3.0,2.5,1.76,1.54,1.33,1.32,1,0.88,0.5,0.22))
#q-qplot first method:
i<-1:length(x)
W<-log(-log(1-(i/(length(x)+1))))
plot(W,log(sort(x)),col="Orange",xlab="Wi",ylab="Ln(ti)")
fit1<-lm(log(sort(x))~W)
abline(fit1,col="Brown")
```

نمودار احتمال توزیع های وایبل و مقدار غایی

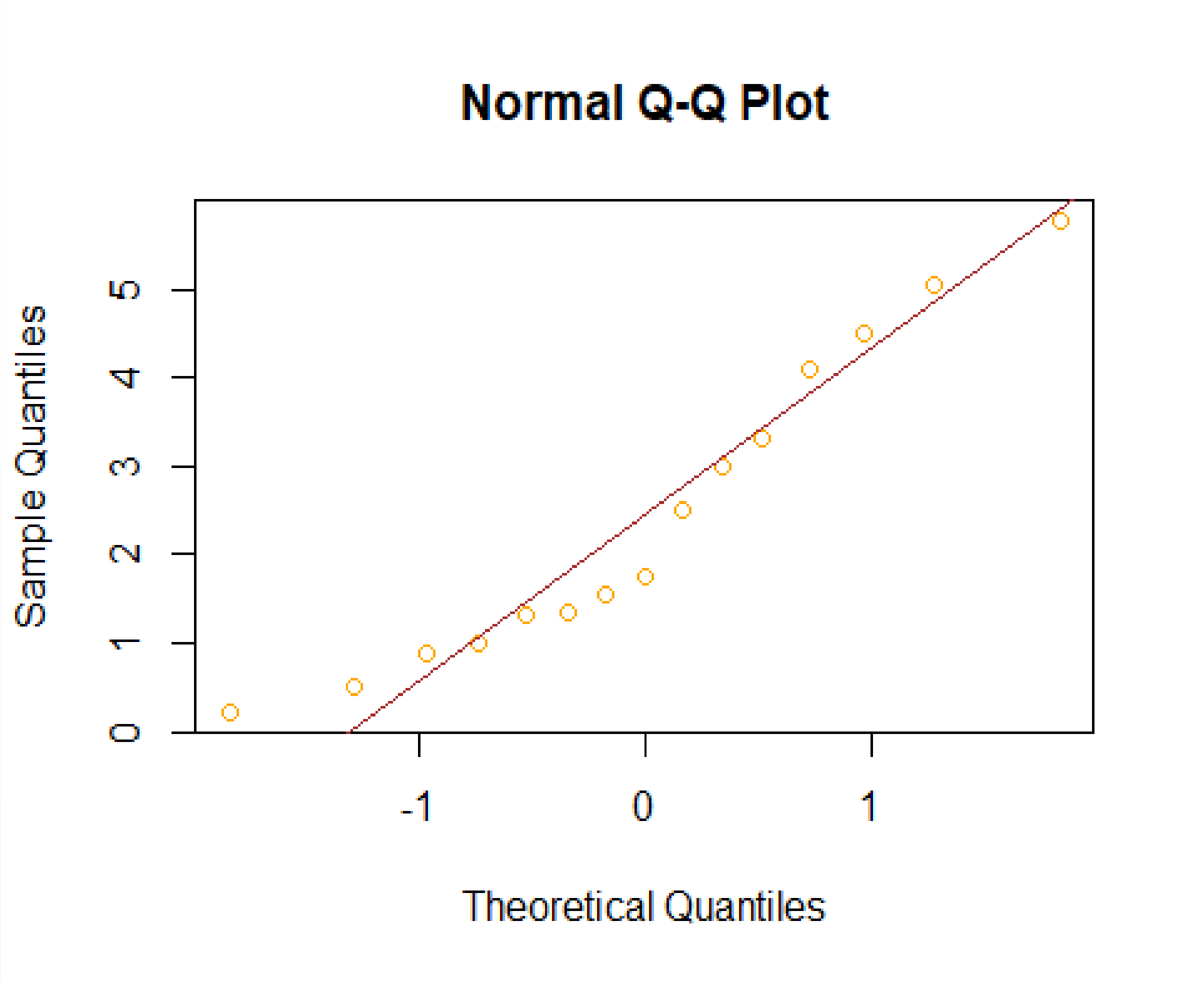



```
fit1$coefficients

## (Intercept)           W
##   1.0233689   0.8623622

lamba=exp(-fit1$coefficients[1]);Betha.hat=1/fit1$coefficients[1]
print(paste("betha.hat is equal to ",Betha.hat,
            "the lambda is equal to",lamba))

## [1] "betha.hat is equal to  0.977164714202085 the lambda is equal to 0.
359382168030245"
```

نمودار احتمال توزیع های وایبل و مقدار غایی

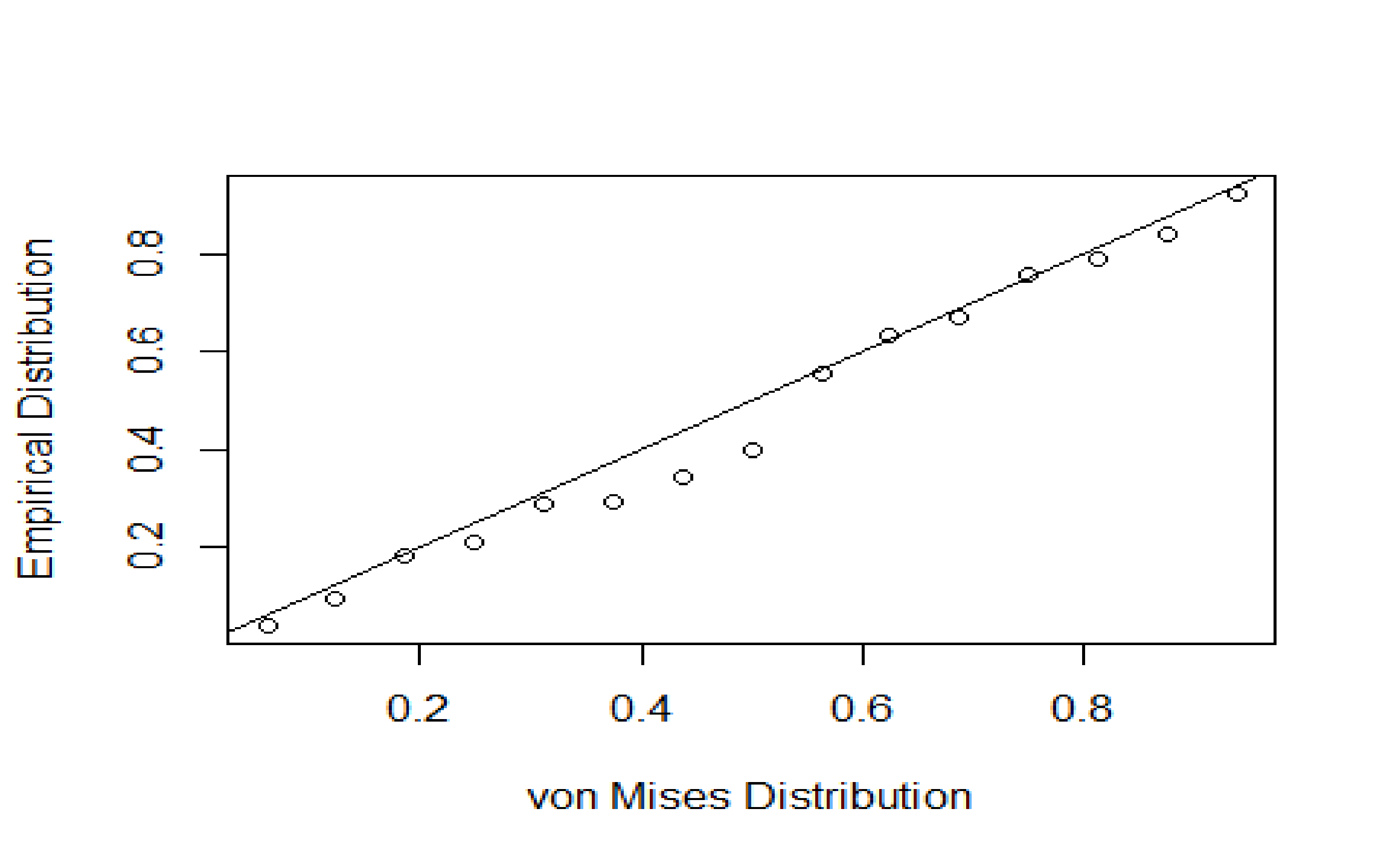
```
# q-qplot second method:
qqnorm(x,col="Orange")
qqline(x,col="brown")
```

نمودار Q-Q



نمودار احتمال توزیع های وایبل و مقدار غایی

```
# p-pplot:
pp.plot(x,ref.line = TRUE)
```




```
##           mu     kappa
## 1 1.303468 0.510308
```

نمودار احتمال توزیع های وایبل و مقدار غایی

خروجی SPSS

Case Processing Summary

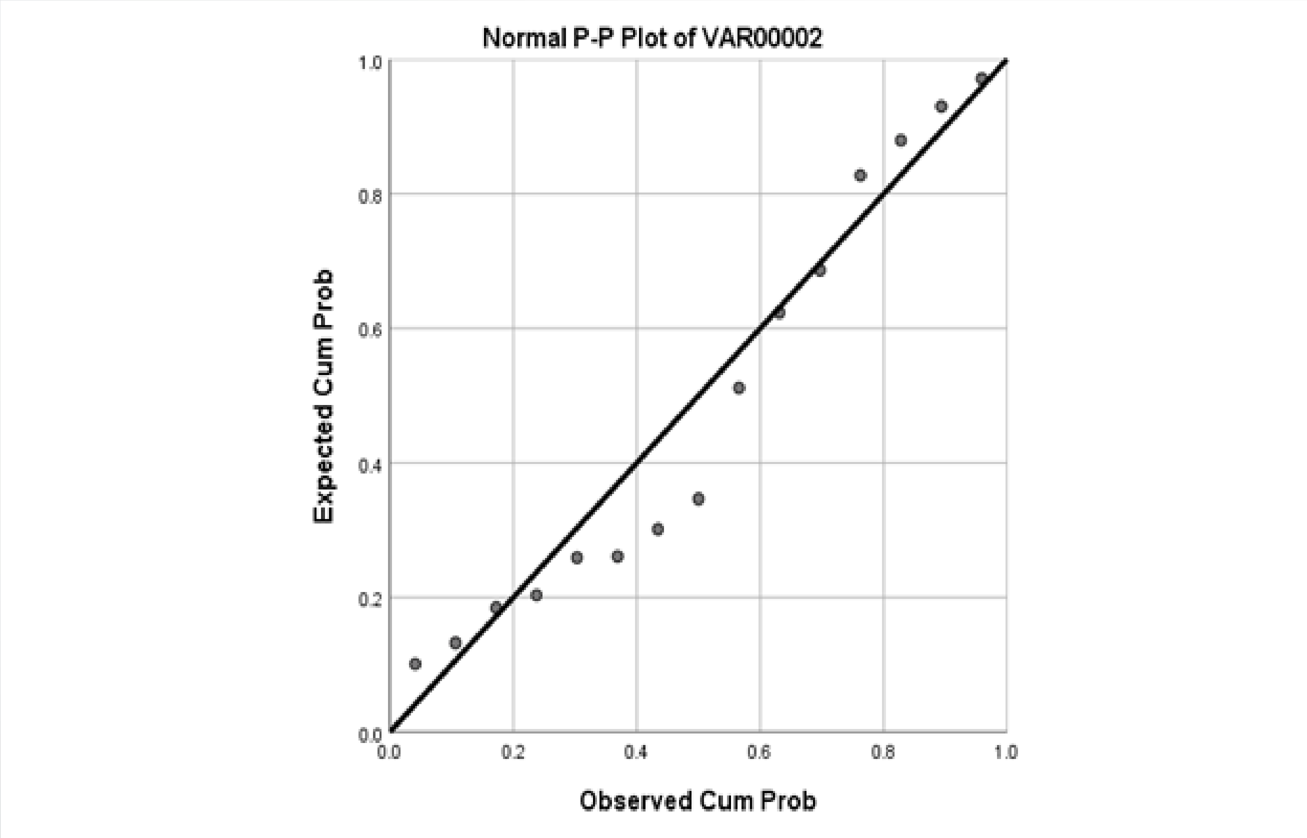
		VAR00002
Series or Sequence Length		22
Number of Missing Values in the Plot	User-Missing	0
	System-Missing	7

Estimated Distribution Parameters

		VAR00002
Normal Distribution	Location	2.4500
	Scale	1.74866

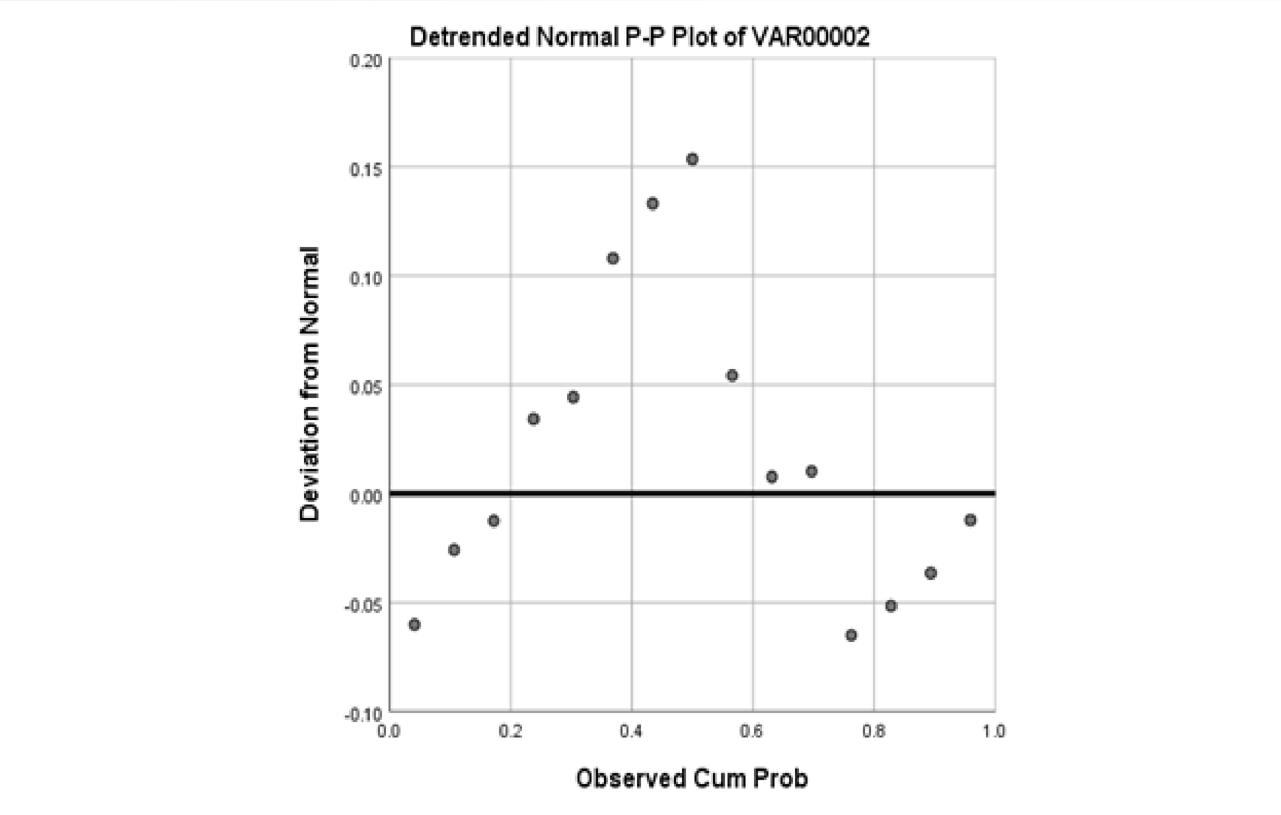
نمودار احتمال توزیع های وایبل و مقدار غایی

نمودار P-P



نمودار احتمال توزیع های وایبل و مقدار غایی

نمودار P-P



نمودار احتمال توزیع های وایبل و مقدار غایی

خروجی SPSS

Case Processing Summary

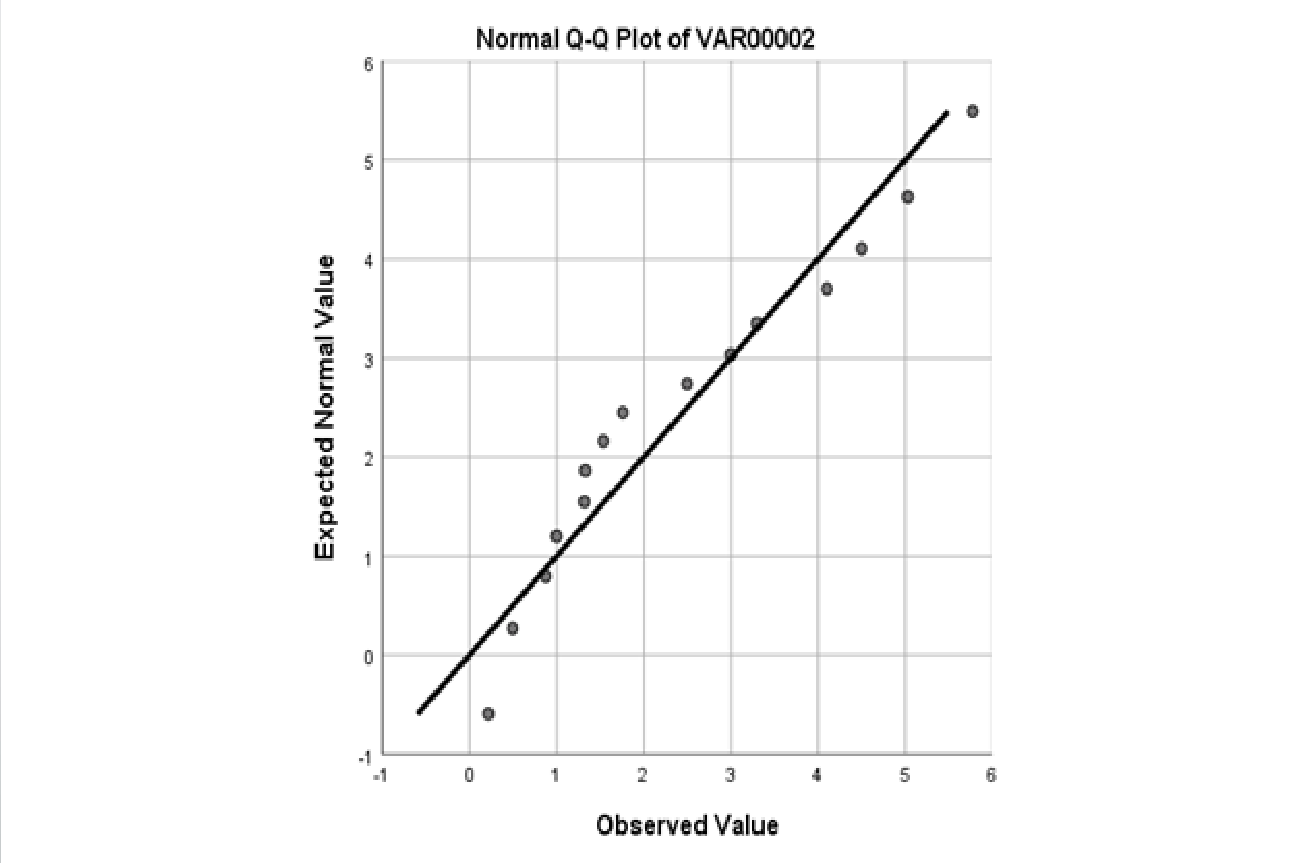
		VAR00002
Series or Sequence Length		22
Number of Missing Values in the Plot	User-Missing	0
	System-Missing	7

Estimated Distribution Parameters

		VAR00002
Normal Distribution	Location	2.4500
	Scale	1.74866

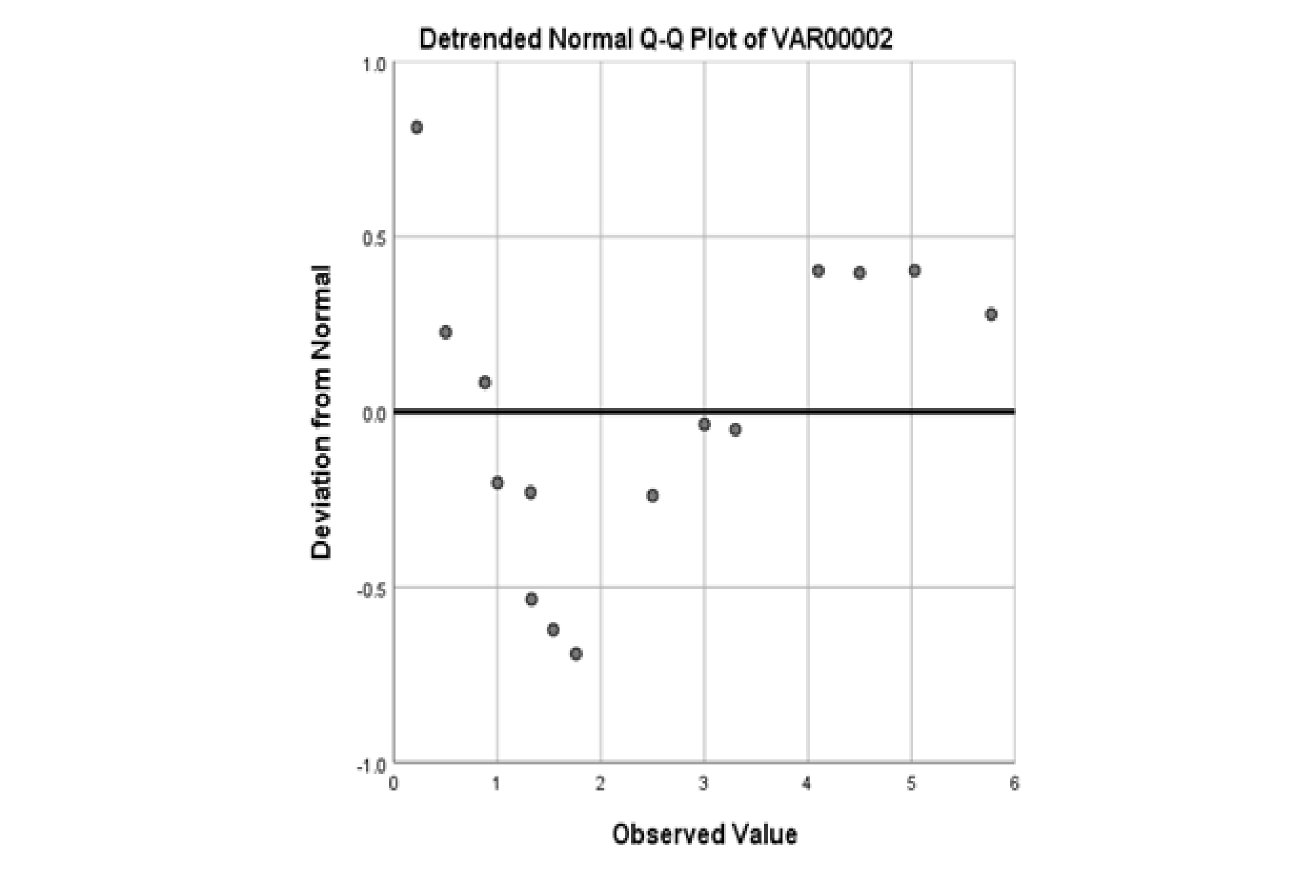
نمودار احتمال توزیع های وایبل و مقدار غایی

نمودار Q-Q



نمودار احتمال توزیع های وایبل و مقدار غایی

نمودار Q-Q



نمودار احتمال توزیع های وایبل و مقدار غایی

اگر توزیع طول عمر توزیع مقدار غایی با تابع توزیع زیر باشد :

$$F(t) = 1 - e^{-e^{\left(\frac{t-\mu}{\sigma}\right)}}, \qquad -\infty < t < \infty$$

آنگاه جهت رسم نمودار احتمال و برآورد پارامترهای μ و σ مراحل زیر را انجام می‌دهیم.

- t_i را در مقابل $W_i = \ln(-\ln(1 - \frac{i}{n+1}))$ رسم می‌کنیم.
- برآورد های به ترتیب عبارت عرض از مبدا و ضریب زاویه خط برازش داده شده خواهند بود.

نمودار احتمال توزیع نرمال و لگ نرمال

فرض کنید توزیع مورد نظر نرمال $N(\mu,\sigma^2)$ باشد. در این حالت مراحل زیر را انجام دهید،

- t_i را در مقابل $Z_i = \Phi^{-1}(\frac{i}{n+1})$ و $i = 1,2,\ldots,n$ رسم کنید که در آن $\Phi(0)$ تابع توزیع نرمال استاندارد است.
- برآورد μ و σ به ترتیب برابر با عرض از مبدا و ضریب زاویه خط برازش داده خواهند بود.

نمودار احتمال توزیع نرمال و لگ نرمال

در صورتی که توزیع طول عمر لگ نرمال با توزیع

$$F(t) = \Phi(\frac{\ln t \ - \ \mu}{\sigma})$$

باشد آنگاه مراحل رسم نمودار احتمال بصورت زیر است

- $\ln t_i$ را در مقابل $Z_i = \Phi^{-1}(\frac{i}{n+1})$ و $i = 1,2, \dots ,n$ رسم می‌کنیم.
- برآورد μ و σ به ترتیب برابر با عرض از مبدا و ضریب زاویه خط برازش داده می‌باشد.

نمودار احتمال توزیع نرمال و لگ نرمال

مثال. داده های زیر زمان شکست مرتب شده ۲۳ بلبرینگ (برحسب میلیون دور) را نشان می‌دهد که در یک آزمایش طول عمر به دست آمده اند.

۱۷/۸۸و۲۸/۹۲و۳۳/۰۰و۴۱/۵۲و۴۱/۱۲و۴۵/۶۰و۴۸/۴۰و۵۱/۸۴و۵۱/۹۵و۵۴/۱۲و۵۵/۵۶و۶۷/۸۰و۶۴/۶۴و

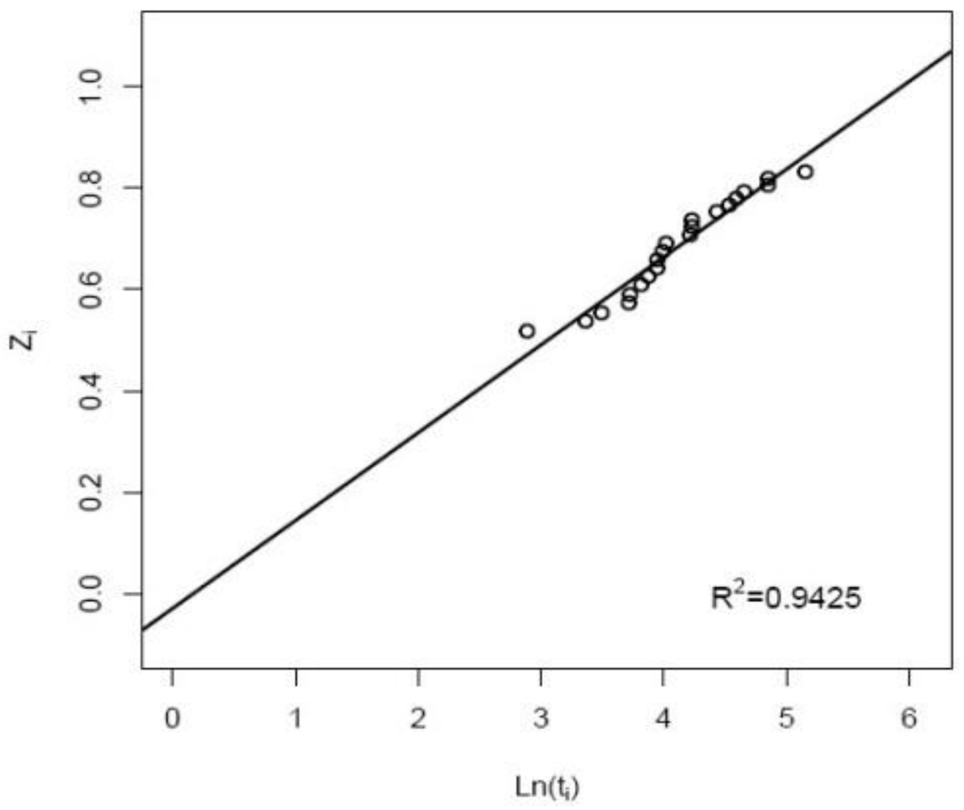
۶۸/۶۴و ۸۴/۱۲و۹۳/۱۲و۹۸/۶۴و۱۰۵/۱۲و۱۲۷/۹۲و۱۲۸/۰۴و۱۷۳/۰۴

آیا می‌توان پذیرفت که توزیع طول عمر بلبرینگ ها لگ نرمال است؟ در صورت مثبت بودن جواب پارامترهای توزیع را برآورد کنید.

نمودار احتمال توزیع نرمال و لگ نرمال

حل. با رسم نقاط $(\ln t_i , z_i)$ و $i = 1,2, \ldots ,23$ ، شکل ۳.۸ به دصست می آید. با توجه به قرار گرفتن نقاط حول خط برازش داده شده و مقدار مربع R میتوان پذیرفت که توزیع طول عمر بلبرینگ ها لگ نرمال است. برآوردهای و σ به ترتیب برابر با عرض از مبدا و ضریب زاویه خط برازش داده یعنی ۰/۳۹۹و ۵/۴۳۳ می‌باشد.

نمودار احتمال توزیع نرمال و لگ نرمال



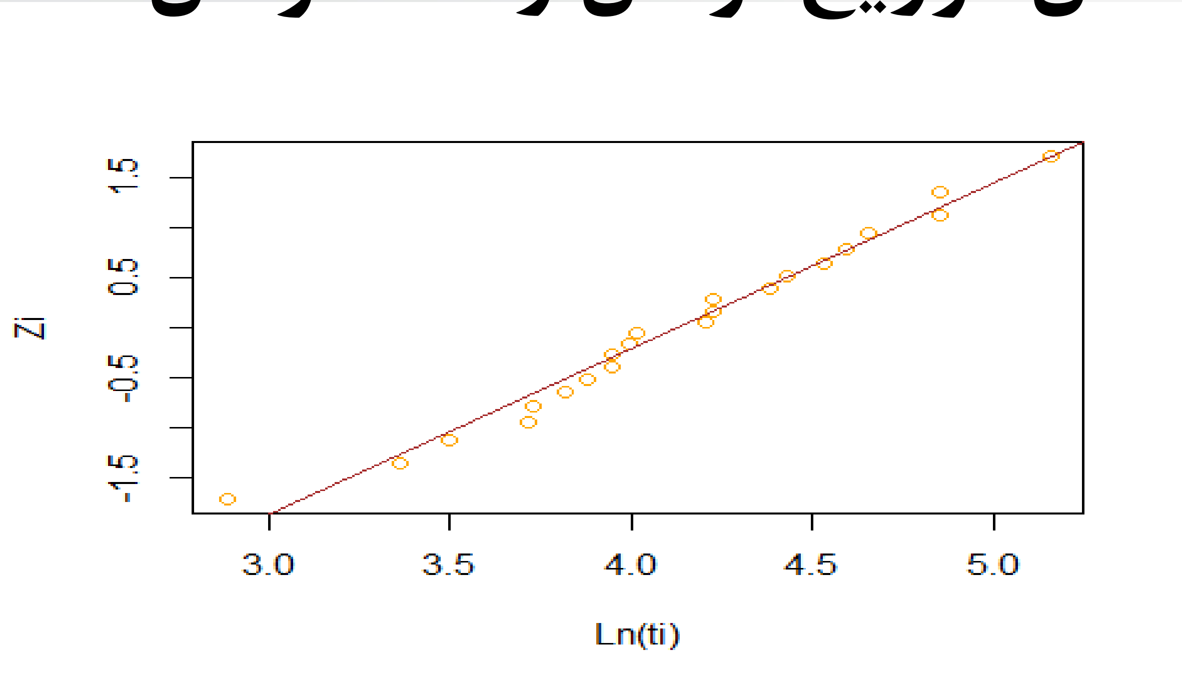
شکل ۳.۸ نمودار احتمال توزیع لگ نرمال برای داده‌های طول عمر بلبرینگ

نمودار احتمال توزیع نرمال و لگ نرمال

خروجی نرم افزار R

```
#example 3.8:
t<-c(17.88,28.92,33.00,41.52,41.12,45.60,48.40,51.84,51.95,54.12,55.56,67,
80,68.64,68.64,84.12,93.12,98.64,105.12,127.92,128.04,173.40)
#q-qplot first method:
i<-1:length(t)
Zi<-qnorm(i/(length(t)+1))
plot(log(sort(t)),Zi,xlab = "Ln(ti)",ylab= "Zi",col="Orange")
fit<-lm(Zi~log(sort(t)))
abline(fit,col="brown")
```

نمودار احتمال توزیع نرمال و لگ نرمال

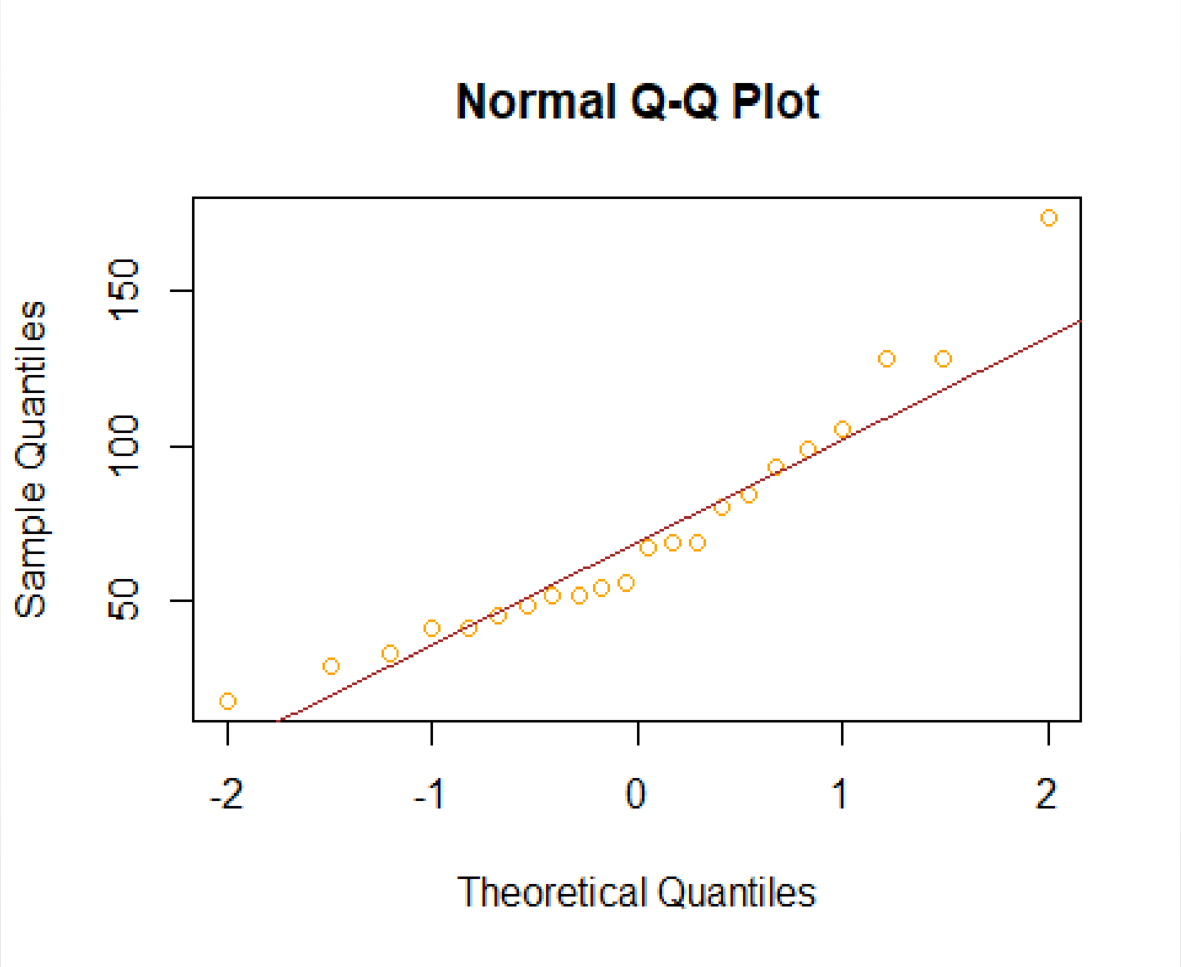



```
fit$coefficients

##   (Intercept) log(sort(t))
##     -6.842195      1.655786

mu.hat<-fit$coefficients[1];sigma.hat<-fit$coefficients[2]
print(paste("the mu.hat is equal to",mu.hat,
            "the sigma.hat is equal to",sigma.hat))

## [1] "the mu.hat is equal to -6.84219457592908 the sigma.hat is eq
ual to 1.65578644318971"
```

```
# q-qplot second method:
qqnorm(t,col="Orange")
qqline(t,col="brown")
```

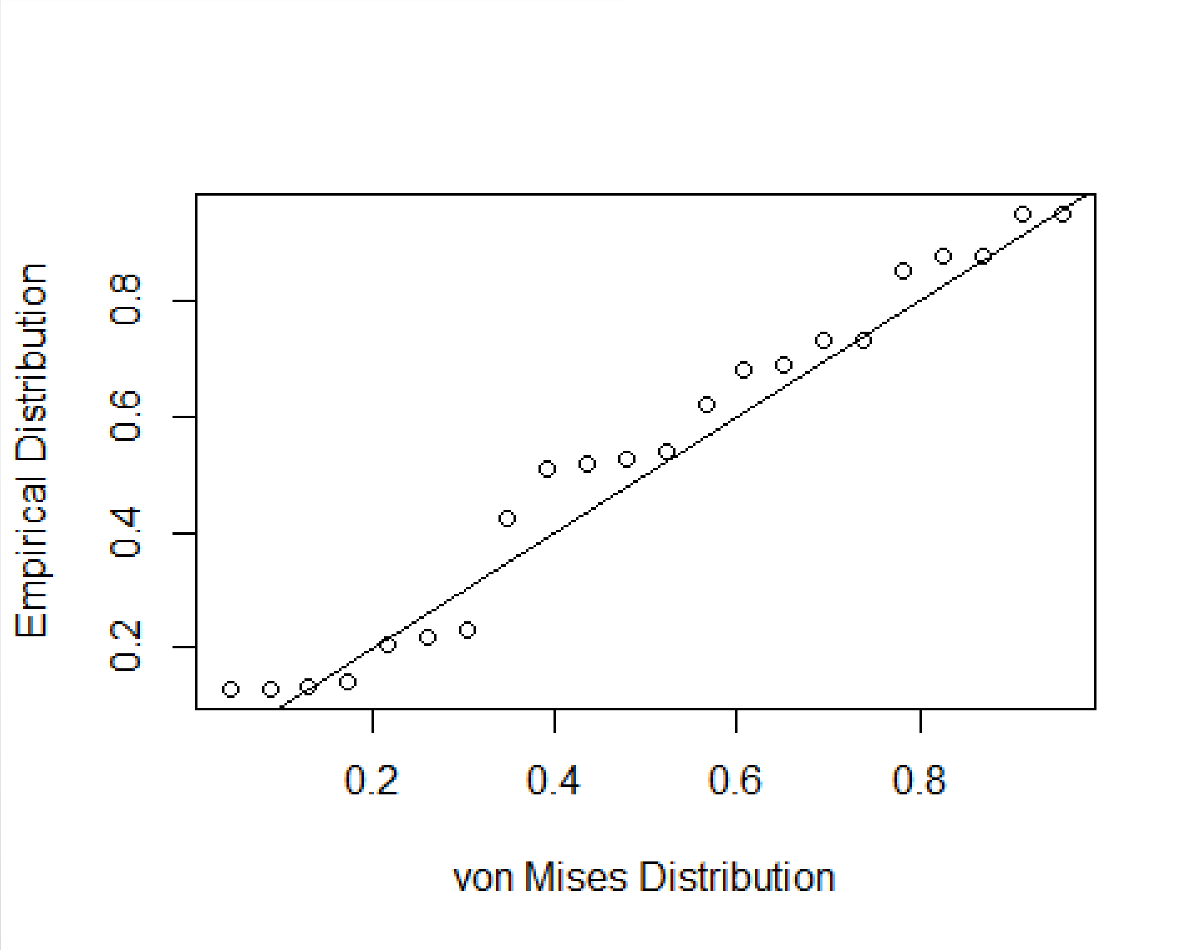
نمودار احتمال توزیع نرمال و لگ نرمال

نمودار Q-Q



نمودار احتمال توزیع نرمال و لگ نرمال

```
#p-pplot :
pp.plot(t,ref.line = TRUE)
```




```
##         mu     kappa
## 1 3.971496 0.6268875
```

نمودار احتمال توزیع نرمال و لگ نرمال

خروجی SPSS

Case Processing Summary

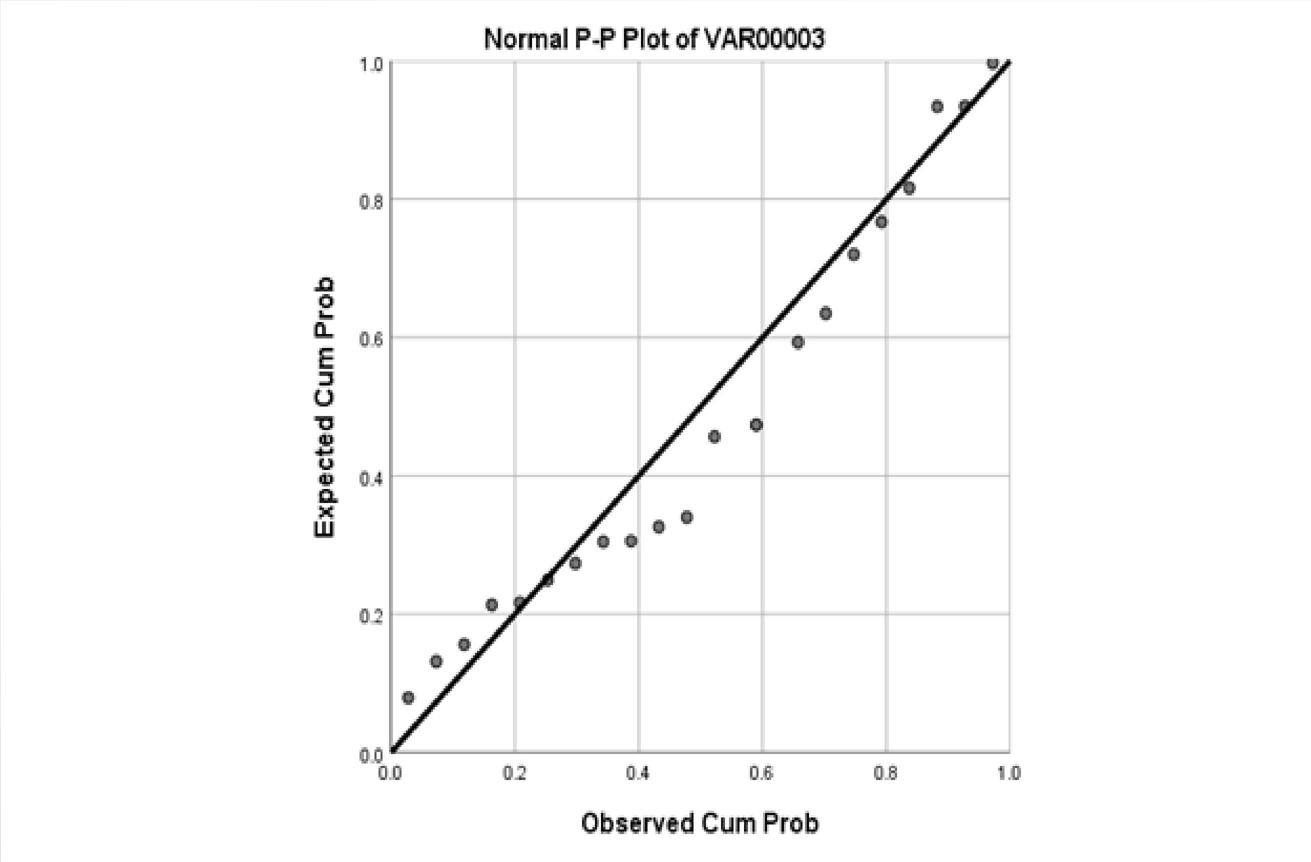
		VAR00003
Series or Sequence Length		22
Number of Missing Values in the Plot	User-Missing	0
	System-Missing	0

Estimated Distribution Parameters

		VAR00003
Normal Distribution	Location	71.1164
	Scale	37.72372

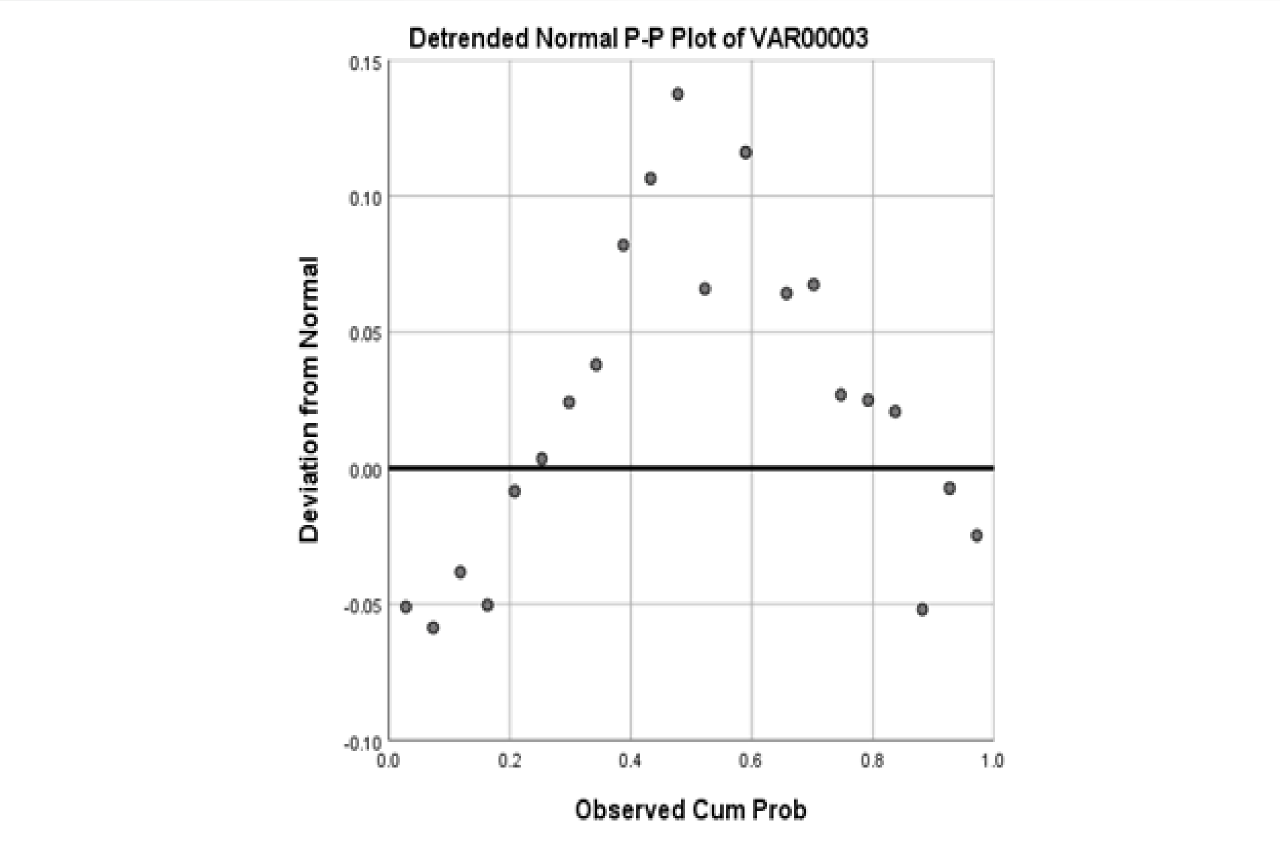
نمودار احتمال توزیع نرمال و لگ نرمال

نمودار P-P



نمودار احتمال توزیع نرمال و لگ نرمال

نمودار P-P



نمودار احتمال توزیع نرمال و لگ نرمال

خروجی SPSS

Case Processing Summary

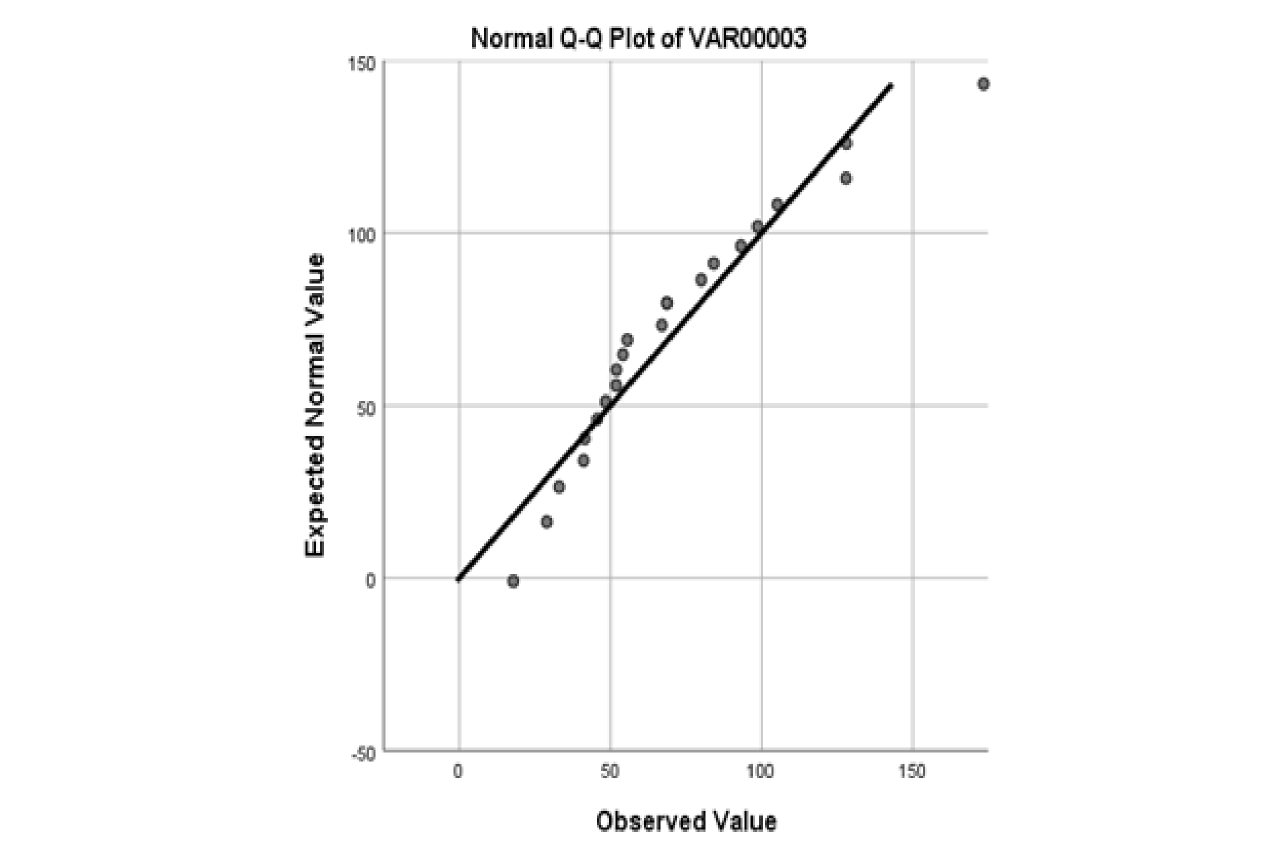
		VAR00003
Series or Sequence Length		22
Number of Missing Values in the Plot	User-Missing	0
	System-Missing	0

Estimated Distribution Parameters

		VAR00003
Normal Distribution	Location	71.1164
	Scale	37.72372

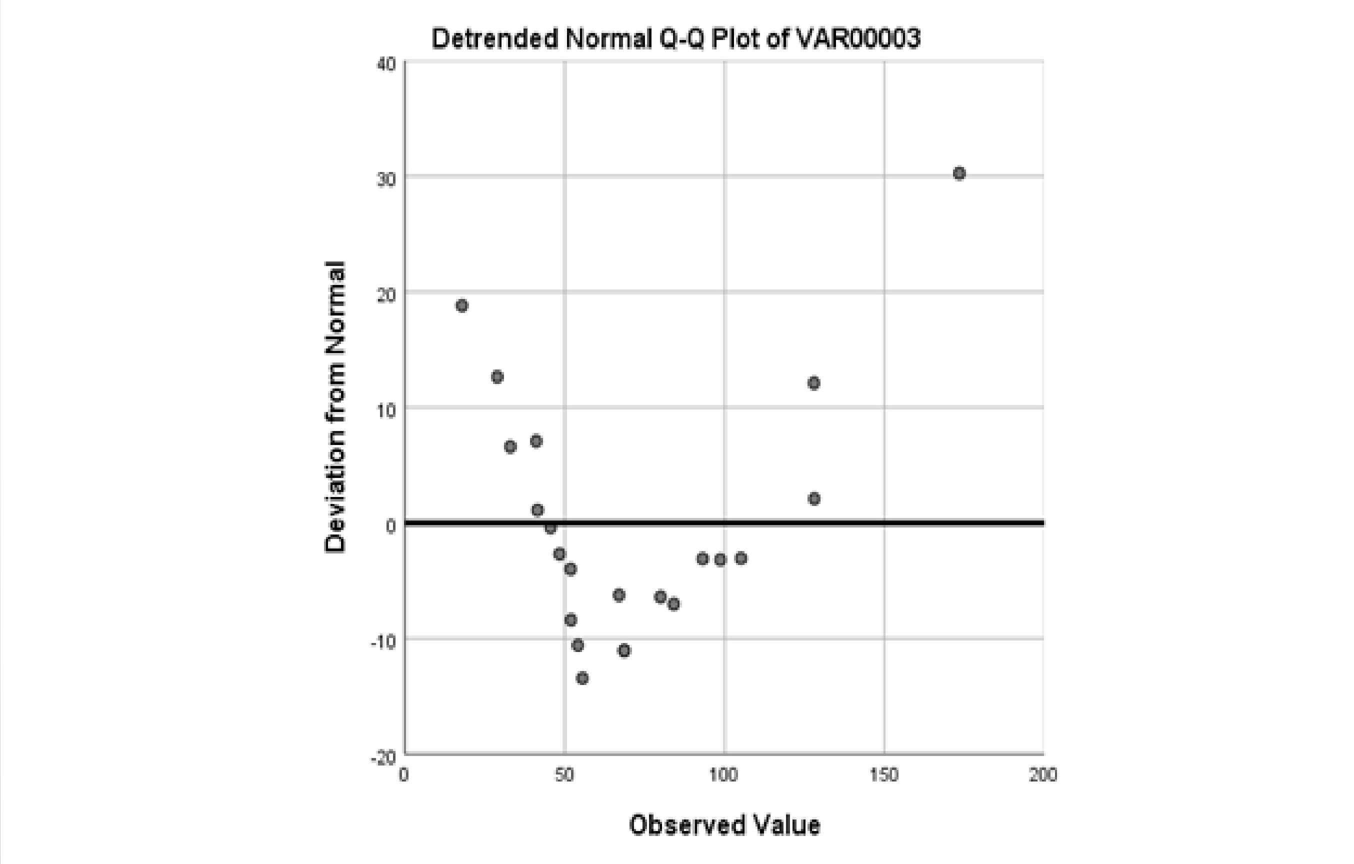
نمودار احتمال توزیع نرمال و لگ نرمال

نمودار Q-Q



نمودار احتمال توزیع نرمال و لگ نرمال

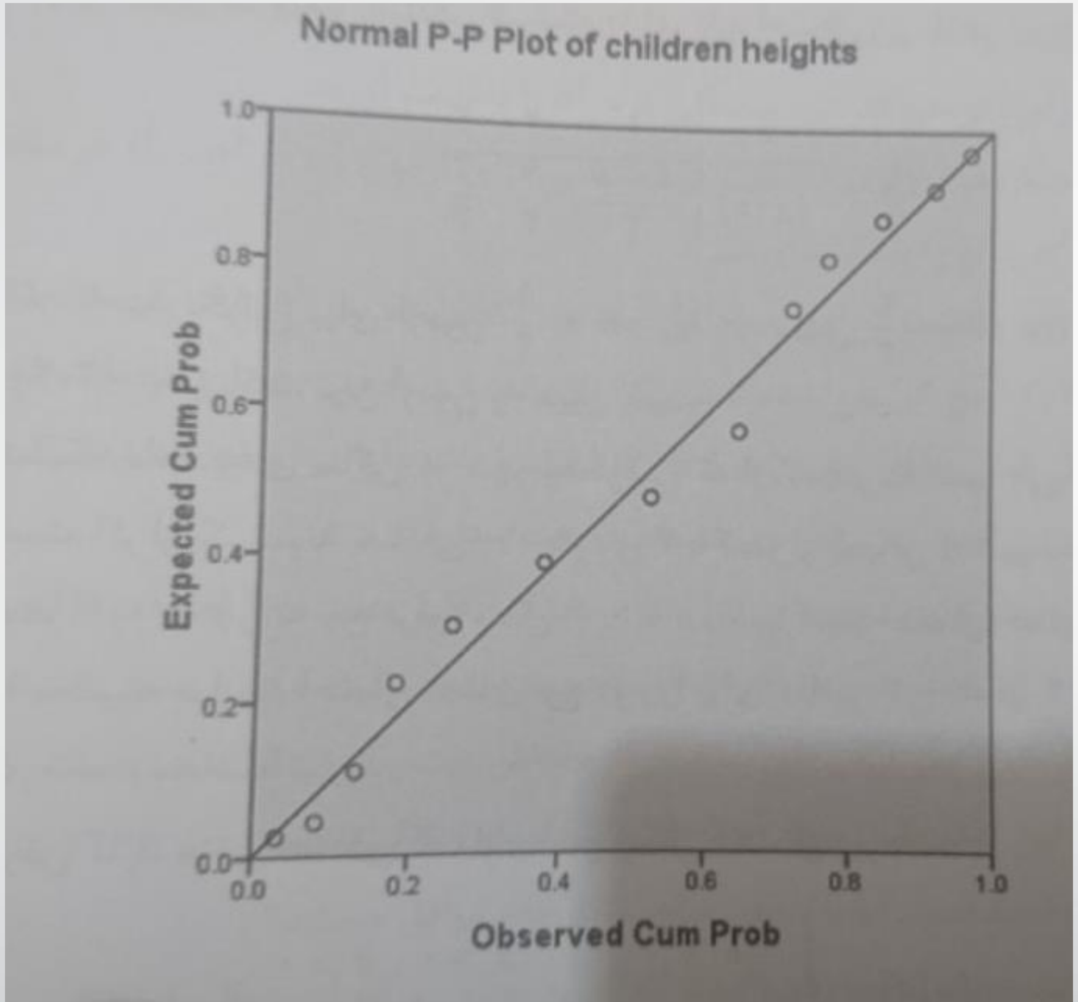
نمودار Q-Q



جدول ۱.۸ خلاصه مطالب ارائه شده در این بخش به همراه مختصات مورد نیاز برای رسم نمودار احتمال و برآورد پارامترهای چند توزیع دیگر را نشان می‌دهد که در آن فرض کرده ایم a ضریب زاویه و b عرض از مبدا خط برازش داده شده به داده ها می‌باشد.

جدول ۱.۸ خلاصه چگونگی رسم نمودار احتمال برای توزیع‌های مختلف

برآورد	پارامتر	مختصات رسم نمودار	توزیع
$\frac{1}{a}$	λ	$\left(t_i, -\ln\left(1-\frac{i}{n+1}\right)\right)$	نمایی $E(\lambda,\theta)$
b	θ		
e^{-b}	λ	$\left(\ln t_i, \ln\left(-\ln\left(1-\frac{i}{n+1}\right)\right)\right)$	وایبل $W(\lambda,\beta)$
$\frac{1}{a}$	β		
b	μ	$\left(t_i, \ln\left(-\ln\left(1-\frac{i}{n+1}\right)\right)\right)$	مقدار غایی $Ex(\mu,\sigma)$
a	σ		
b	μ	$\left(t_i, \Phi^{-1}\left(\frac{i}{n+1}\right)\right)$	نرمال $N(\mu,\sigma^2)$
a	σ		
b	μ	$\left(\ln t_i, \Phi^{-1}\left(\frac{i}{n+1}\right)\right)$	لگ نرمال $LN(\mu,\sigma)$
a	σ		
$\frac{1}{a}$	λ	$\left(\ln t_i, \Phi^{-1}\left(\frac{i}{n+1}\right)\right)$	گاما $G(\alpha,\lambda)$ (با فرض معلوم بودن α)
a	μ	$\left(t_i, \ln\left(\frac{i}{n+1-i}\right)\right)$	لجستیک (μ,σ)
b	σ		
a	μ	$\left(\ln t_i, \ln\left(\frac{i}{n+1-i}\right)\right)$	لگ لجستیک (μ,σ)
b	σ		



شکل ۹.۴: نمودار $P-P$ مربوط به قد کودکان ۱۰ ساله

نمودار احتمال نرمال داده های سانسور شده

روش رسم نمودار های احتمال برای داده های سانسور شده از راست ، مثلا داده های سانسور شده نوع I,II، کاملا مشابه داده های کامل است. در این حالت، نمودار احتمال را برای داده هایی که زمان شکست آنها داده شده است ،رسم می‌کنیم. یعنی $t(i)$ را در مقابل $F^{-1}(\frac{i}{n+1})$ رسم می‌کنیم که در آن i=1,2,3,…,r تعداد شکست ها و n تعداد کل مشاهدات است. برای مثال، اگر n=20 واحد در آزمایش قرار گیرد و زمان شکست ۱۵ واحد اول را مشاهده کنیم، آنگاه موقعیت های نمودار به صورت جدول ۳ . ۷ خواهد بود.

جدول ۳.۷ موقعیت‌های نمودار در حالت سانسور از سمت راست

i	۱	۲	...	۱۴	۱۵	۱۶	...	۲۰
$t_{(i)}$	$t_۱$	$t_۲$	...	$t_{۱۴}$	$t_{۱۵}$	-	...	-
$\frac{i}{n+۱}$	$\frac{۱}{۲۱}$	$\frac{۲}{۲۱}$	...	$\frac{۱۴}{۲۱}$	$\frac{۱۵}{۲۱}$	-	...	-

در اینجا فقط ۱۵ نقطه $(t(i) ,F^{-1}(\frac{i}{21}))$ و $i = 1,2, \dots ,15$ را رسم می‌کنیم. اگر داده ها از چپ (یا هم از چپ و هم از راست) سانسور شوند نیز، روش رسم نمودار مشابه است. یعنی فقط زمان های شکست را در مقابل $F^{-1}(\frac{i}{n+1})$ رسم می‌کنیم که در آن i اندیس زمان های شکست است.

در رسم نمودار روی کاغذ های احتمال نیز ، روش رسم نمودار به همین صورت است یعنی زمان های شکست را در مقابل $MR\% = \frac{i-0/3}{n-0/4} * 100$ رسم می‌کنیم که در آن i زمان های شکست، و n تعداد کل واحد هایی است که در آزمایش گرفته اند.

برای توضیح بیشتر مثال بعدی را در نظر می‌گیریم:

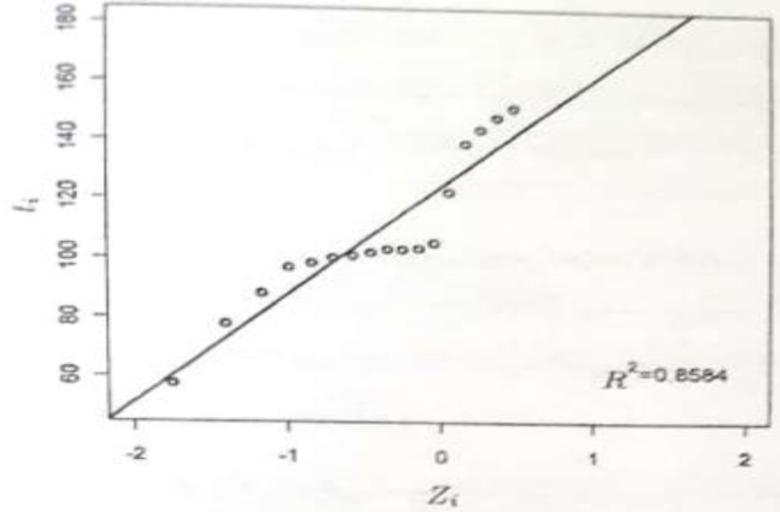
مثال. در یک نوع وسیله الکتریکی نوعی سیم بکار می‌رود. برای اندازه گیری مقاومت سیم، آن را با یک سیستم خم کن مورد آزمایش قرار می‌دهند. اطلاعات مربوط به زمان شکست داده ها (بر حسب ساعت) در جدوا ۷ . ۴ داده شده است. آیا می‌توان پذیرفت که داده ها نرمال هستند؟ در صورت مثبت بودن جواب ، پارامتر های توزیع را برآورد کنید.

حل. با رسم نقاط $(t(i),z(i))$ ، i=1,2,…,17 شکل ۷ . ۵ به دست می‌آید. باتوجه به نحوه قرار گرفتن نقاط حول خط برازش داده شده و مقدار مربع R ، بطور تقریبی می‌توان پذیرفت که توزیع داده ها نرمال است. با توجه به مطالب بیان شده در بخش پیشین ، برآورد های μ و σ به ترتیب با عرض از مبدا و ضریب زاویه خط برازش داده شده برابر است. پس در این مثال $\hat{\mu} = 124/5$ و $\hat{\sigma} = 36/8$ برآورد می‌شود.

جدول ۴.۷ داده‌های مقاومت سیم

i	۱	۲	۳	۴	۵	۶	۷	۸
t_i	۵۷٫۵	۷۷٫۸	۸۸٫۰	۹۶٫۹	۹۸٫۴	۱۰۰٫۳	۱۰۰٫۸	۱۰۲٫۱
i	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶
t_i	۱۰۳٫۳	۱۰۳٫۳	۱۰۳٫۴	۱۰۵٫۴	۱۲۲٫۶	۱۳۹٫۳	۱۴۳٫۹	۱۴۸٫۰
i	۱۷	۱۸*	۱۹*	۲۰*	۲۱*	۲۲*	۲۳*	۲۴*
t_i	۱۵۱٫۳	۱۶۱٫۱	۱۶۱٫۲	۱۶۱٫۲	۱۶۱٫۴	۱۶۲٫۷	۱۶۳٫۱	۱۷۶٫۸

علامت * در جدول نشان‌دهندهٔ آن است که واحد مورد آزمایش پیش از شکست از آزمایش خارج شده است.



شکل ۵.۷ نمودار احتمال توزیع نرمال برای داده‌های مقاومت سیم.