



**Národní informační středisko
pro podporu jakosti**



OVĚŘOVÁNÍ PŘEDPOKLADU NORMALITY

Doc. Ing. Eva Jarošová, CSc.

Ing. Jan Král



Používané metody

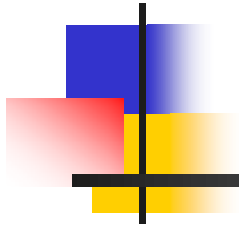
- n statistické testy:
 - n Chí-kvadrát test dobré shody
 - n Kolmogorov - Smirnov
 - n Shapiro - Wilk
 - n Anderson - Darling
 - n Ryan - Joiner
- n grafické metody:
 - n Histogram
 - n Pravděpodobnostní graf
 - n Q - Q graf
 - n P - P graf



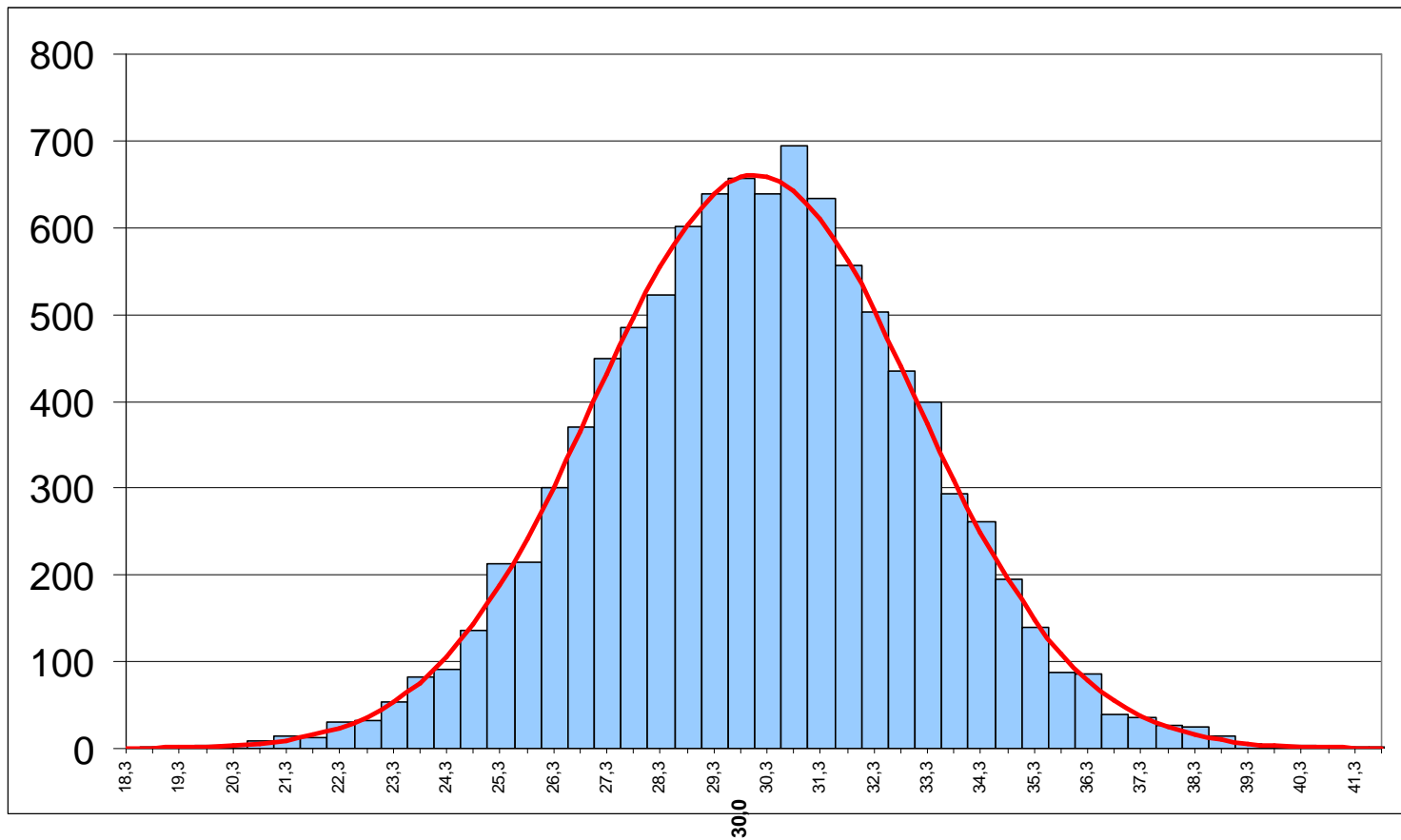
Histogram

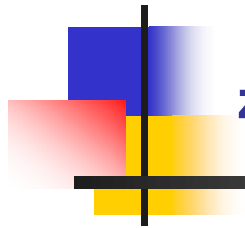
Histogram sestavený na základě dostatečného počtu hodnot pocházejících z normálního rozdělení má charakteristický tvar, jehož modelem je Gaussova křivka.

Příklad histogramu sestaveného z 10 000 hodnot z normálního rozdělení se střední hodnotou $\mu=30$ a směrodatnou odchylkou $\sigma=3$ je na následujícím obrázku.

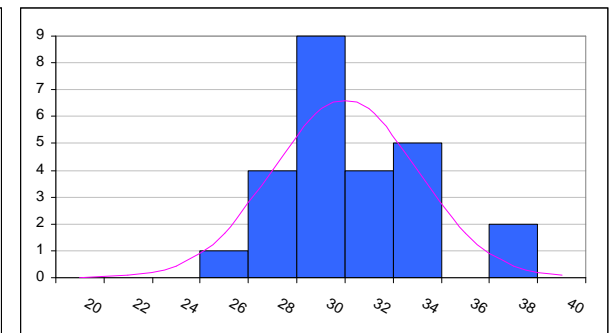
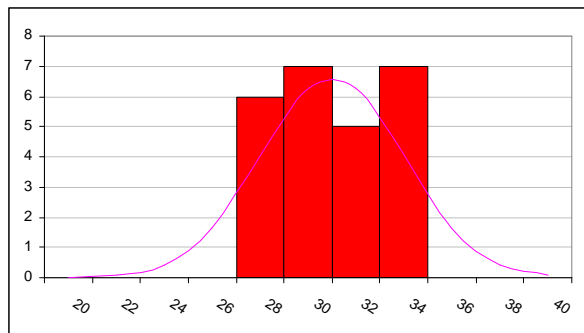
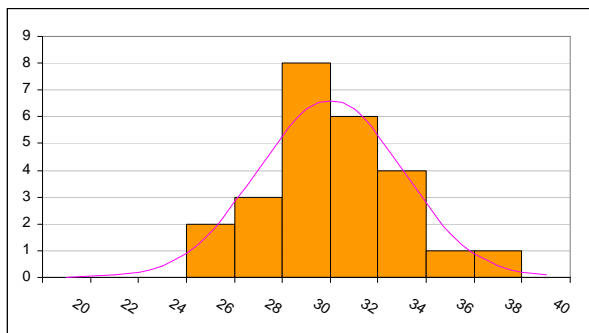
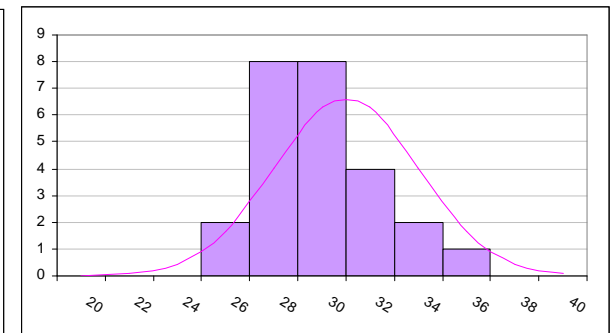
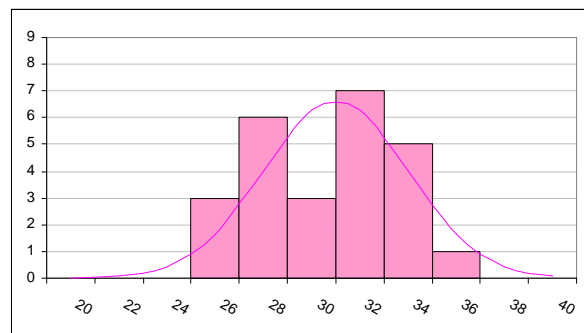
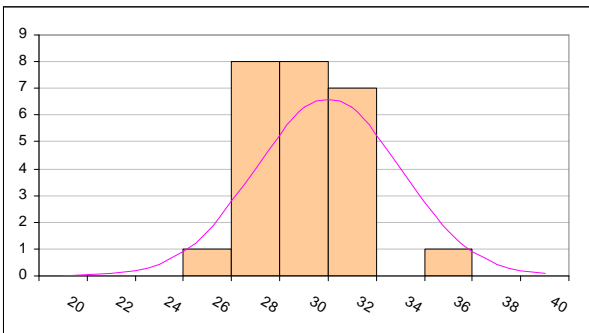
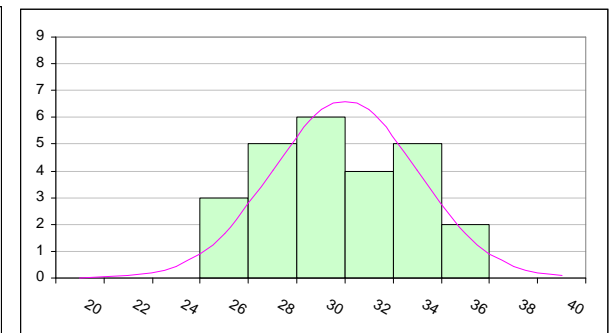
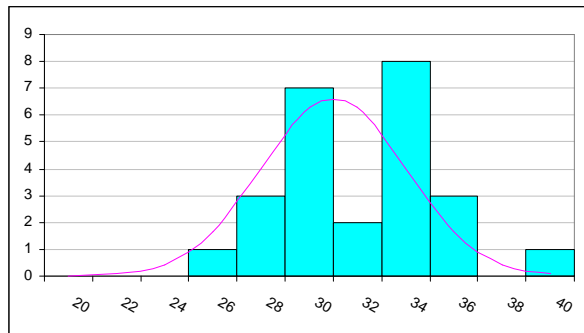
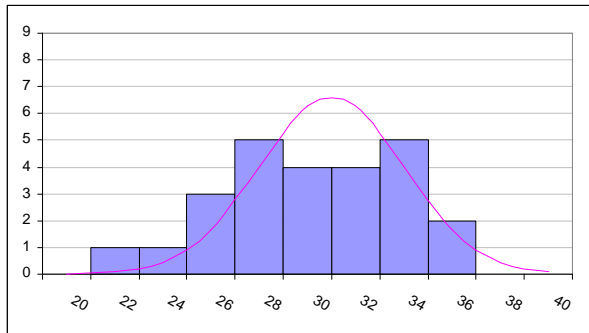


Histogram

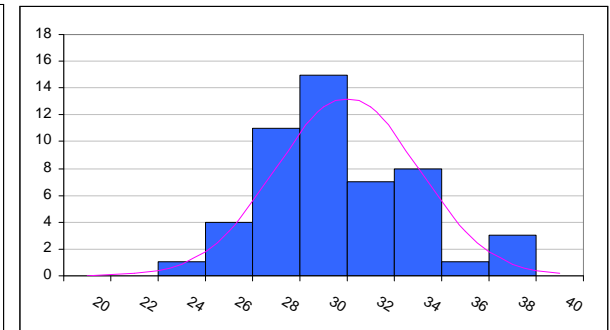
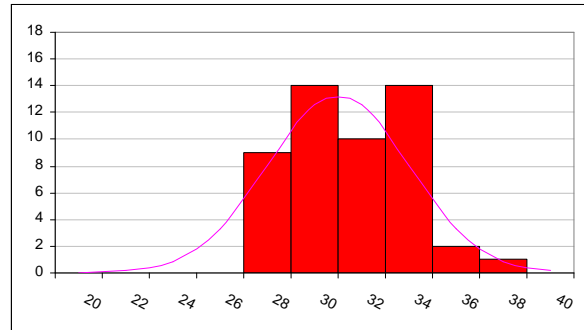
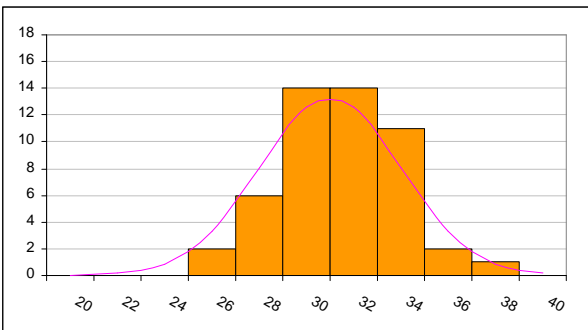
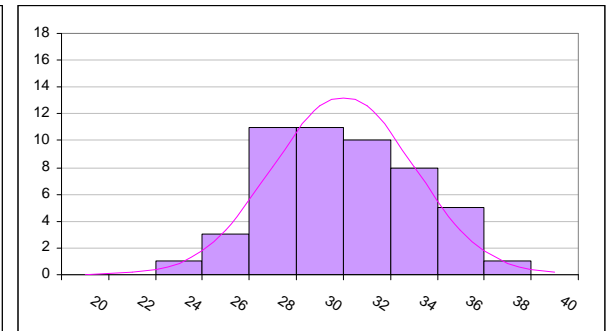
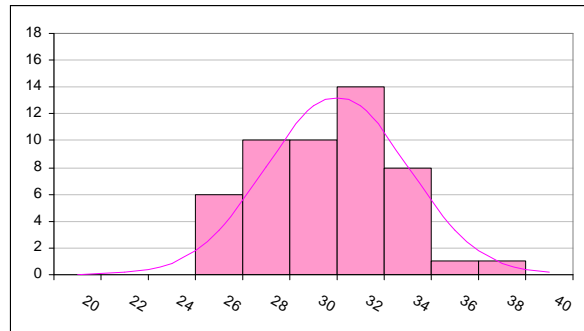
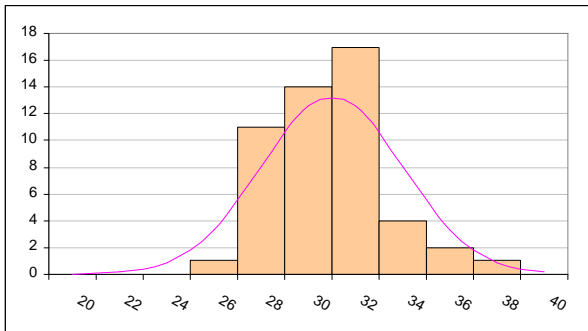
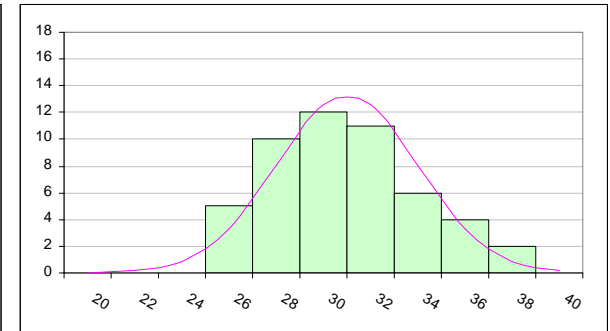
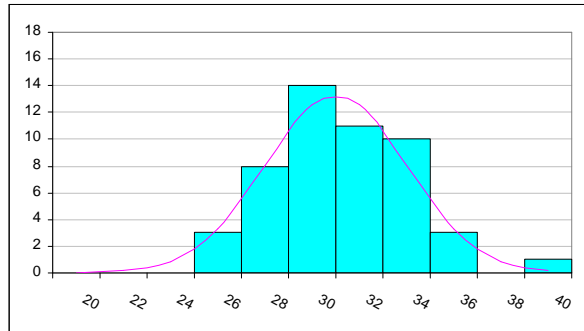
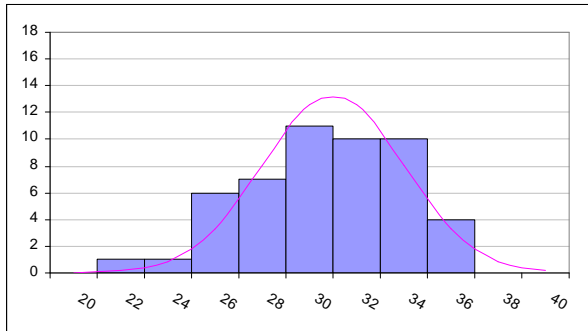




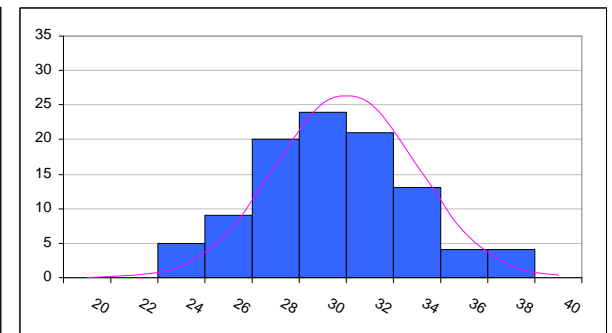
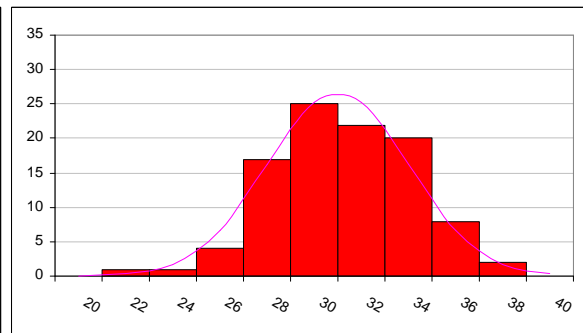
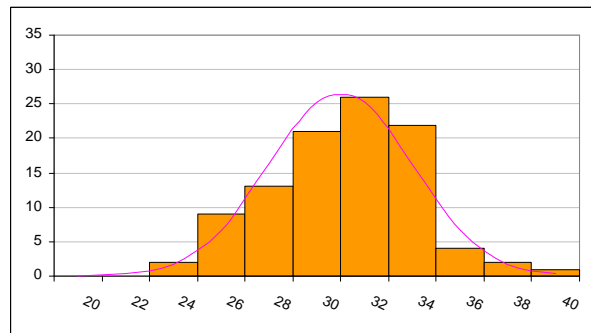
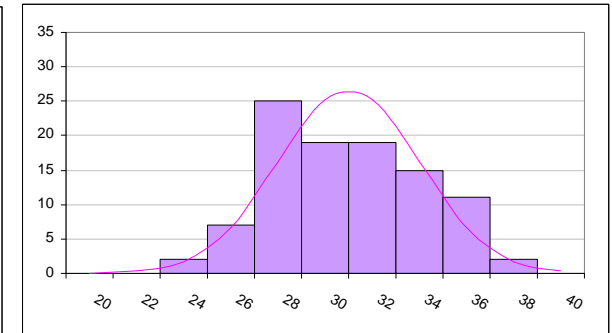
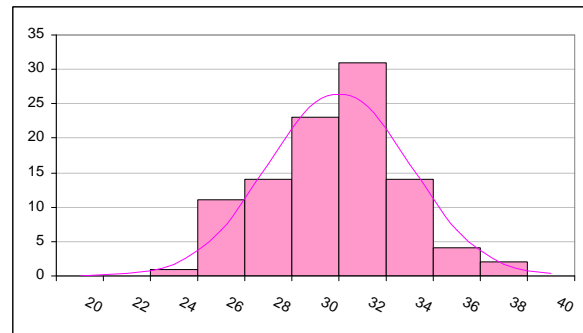
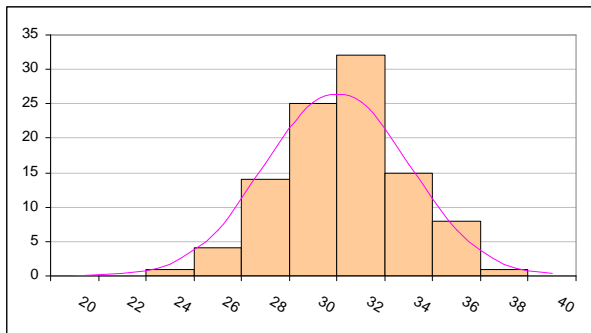
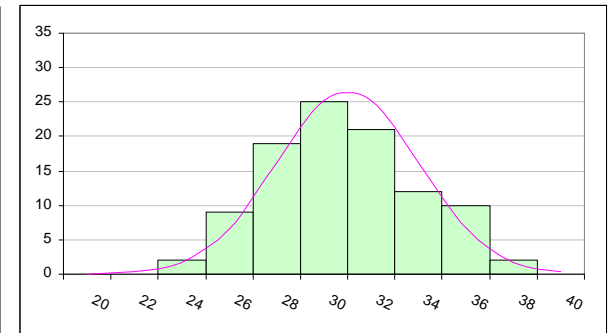
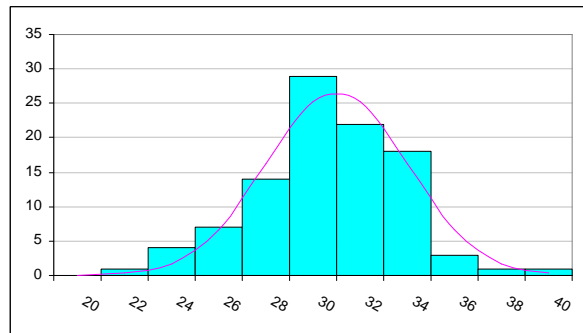
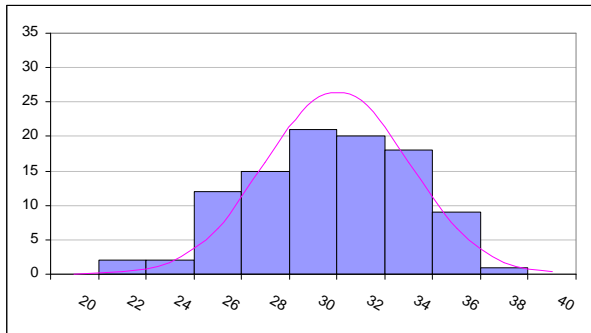
Výběry rozsahu $n = 25$ ze základního souboru s normálním rozdělením $m = 30$ a $s = 3$



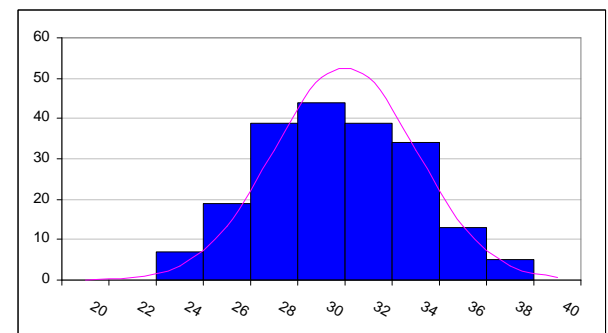
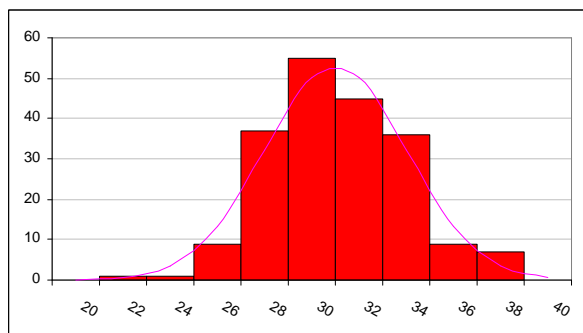
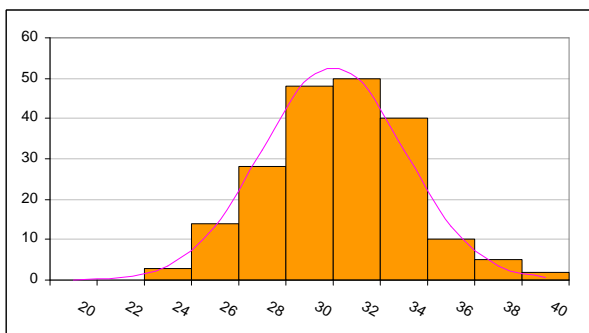
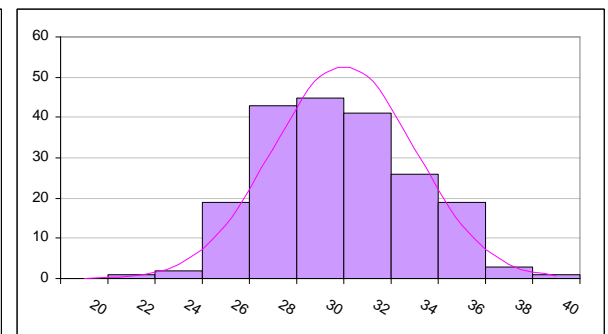
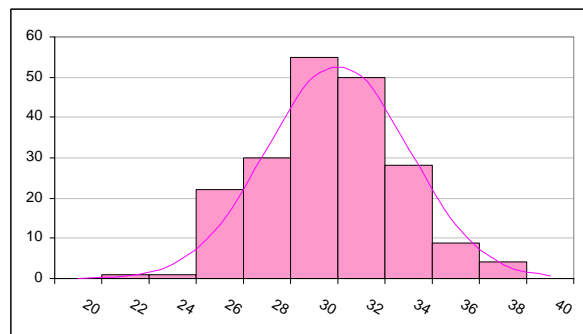
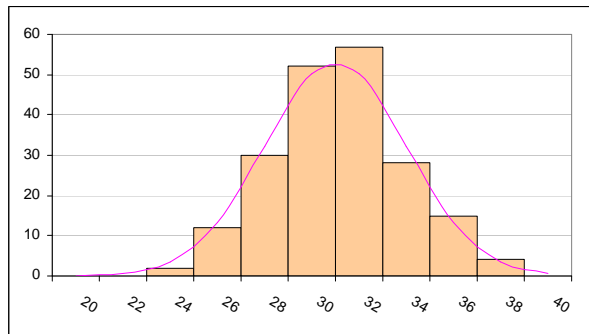
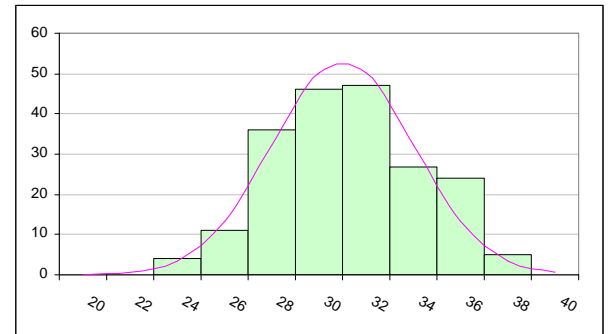
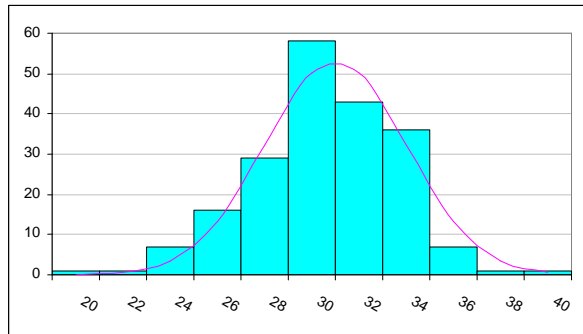
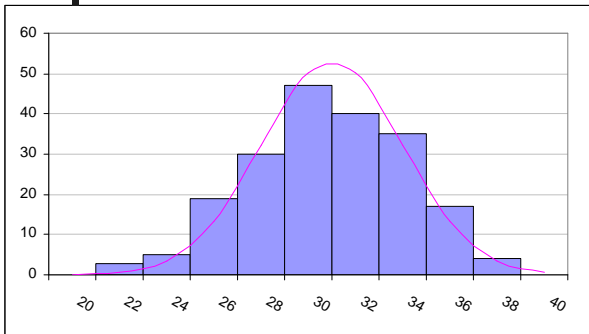
Výběry rozsahu $n = 50$ ze základního souboru s normálním rozdělením $m = 30$ a $s = 3$



Výběry rozsahu $n = 100$ ze základního souboru s normálním rozdělením $m = 30$ a $s = 3$



Výběry rozsahu $n = 200$ ze základního souboru s normálním rozdělením $m = 30$ a $s = 3$





Zhodnocení

Všechny uvedené histogramy představují náhodné výběry z normálního rozdělení se střední hodnotou $\mu = 30$ a směrodatnou odchylkou $\sigma = 3$.

Vidíme, že čím je větší rozsah výběru n , tím lépe odpovídá výběrové rozdělení, znázorněné histogramem, rozdělení v základním souboru, znázorněnému hustotou pravděpodobnosti.

Při běžně používaném rozsahu $n=100$ nemusí být vizuální posouzení objektivní a tvar histogramu může být navíc ovlivněn volbou mezí intervalu.



Testy dobré shody

Pomocí **testů dobré shody** objektivně posoudíme, zda je možno považovat předpoklad normálního rozdělení za splněný.

Testovaná hypotéza

H_0 : Náhodný výběr **pochází** ze základního souboru s normálním rozdělením

Rozlišují se dva případy:

- a) Model normálního rozdělení je plně specifikován, tj. jsou dány střední hodnota μ a rozptyl σ^2 .
- b) Model normálního rozdělení není plně specifikován, střední hodnota a rozptyl se odhadnou z výběrových hodnot.

Rozdíl mezi plně a neúplně specifikovaným modelem se projeví na rozdělení testové statistiky a tedy při rozhodování o tom, zda vypočtená hodnota testové statistiky je či není v kritickém oboru.

Alternativní hypotéza

- a) H_1 : náhodný výběr **nepochází** ze základního souboru s normálním rozdělením s danými parametry μ a σ .
- b) H_1 : náhodný výběr **nepochází** ze základního souboru s normálním rozdělením



Chí - kvadrát test

Náhodný výběr rozsahu n je rozdělen do k intervalů s četnostmi n_j ($j = 1, 2, \dots, k$), horní meze intervalů označíme x_j .

- § Vypočteme teoretické třídní četnosti za předpokladu, že výběr pochází ze základního souboru s normálním rozdělením $N(\mu, \sigma^2)$:
- § Horní meze x_j třídních intervalů převedeme na hodnoty normované proměnné

$$u_j = \frac{x_j - m}{S} ,$$

- § Není-li model plně specifikován, použijeme místo parametru μ výběrový průměr \bar{x} a místo parametru σ výběrovou směrodatnou odchylku s ;
- § Pro každé j vyhledáme odpovídající hodnoty distribuční funkce normovaného normálního rozdělení $\Phi(u_j)$;

§ Určíme teoretické relativní a absolutní třídň četnosti

$$\pi_j = \phi(u_j) - \phi(u_{j-1}) \quad \text{a} \quad n \pi_j ;$$

§ Intervaly, jejichž teoretická absolutní četnost

$n \cdot \pi_j \leq 5$ sloučíme se sousedními intervaly tak, aby byla splněna podmínka $n \cdot \pi_j > 5$

§ Pro redukovaný počet tříd k° vypočteme výrazy

$$\frac{(n_j - np_j)^2}{np_j} ;$$

§ Jejich součtem (přes redukovaný počet tříd k°) dostaneme hodnotu **testové statistiky**

$$c^2 = \sum_{j=1}^{k^\circ} \frac{(n_j - np_j)^2}{np_j}$$

§ **Kritický obor** pro test normality, na hladině významnosti α , je

$$c^2 > c_{1-\alpha}^2 (k^0 - c - 1)$$

kde $c_{1-\alpha}^2 (k^0 - c - 1)$ je $(1-\alpha)$ - kvantil rozdělení C^2
pro $n = k^0 - c - 1$ stupňů volnosti, c je počet odhadovaných parametrů

§ U plně specifikovaného modelu je $c = 0$.

§ Ověřujeme-li jen tvar normálního rozdělení (neúplně specifikovaný model) a parametry μ a σ^2 odhadujeme z výběrových hodnot, je $c = 2$.



PŘÍKLAD 1

V následující tabulce je demonstrován postup výpočtu testové charakteristiky χ^2 pro náhodný výběr rozsahu $n = 100$, ve kterém pozorované hodnoty byly roztrženy do $k = 8$ intervalů. První interval je $(-\infty; 3,94)$, dalších 6 intervalů má šířku $h = 0,02$ a poslední interval je $(4,06; \infty)$. Ze 100 hodnot byl určen výběrový průměr $\bar{x} = 3,999$ a výběrová směrodatná odchylka $s = 0,030$.

Vzhledem k tomu, že krajní intervaly nesplňují požadavek $n\pi_j \geq 5$, sloučíme je se sousedními intervaly. Redukovaný počet tříd je $k^\circ = 6$. Pro počet stupňů volnosti $\nu = k^\circ - 3 = 3$ a pro hladinu významnosti $\alpha = 0,05$ je kritická hodnota $\chi^2_{0,95}(3) = 7,815$.

Jelikož vypočtená hodnota testové charakteristiky $\chi^2 = 1,477$ nespadá do kritického oboru (není větší než kritická hodnota 7,815), nemáme důvod zamítnout hypotézu o tom, že výběr pochází z normálního rozdělení.

Schéma výpočtu testové statistiky chí-kvadrát

horní mez třídního intervalu	třídní četnosti n_j	u_j	$\Phi(u_j)$	π_j	$n \cdot \pi_j$	$n \cdot \pi_j$	n_j	$\frac{(n \cdot \pi_j - n_j)^2}{n \cdot \pi_j}$
3,940	2	-1,93130	0,02672	0,02672	2,67232			
3,960	9	-1,27529	0,10110	0,07438	7,43808	10,11040	11	0,07827
3,980	20	-0,61928	0,26787	0,16676	16,67627	16,67627	20	0,66245
4,000	23	0,03673	0,51465	0,24678	24,67837	24,67837	23	0,11415
4,020	21	0,69274	0,75576	0,24111	24,11133	24,11133	21	0,40149
4,040	17	1,34875	0,91129	0,15553	15,55274	15,55274	17	0,13467
4,060	5	2,00476	0,97751	0,06621	6,62145	8,87089	8	0,08550
4,080	3			0,02249	2,24944			
							$\chi^2 =$	1,47653



Kolmogorovův-Smirnovův test dobré shody

H_0 : náhodný výběr rozsahu n pochází ze základního souboru s normálním rozdělením $N(\mu, \sigma^2)$ s distribuční funkcí $F(x)$ (plně specifikovaný model)

Uvažujeme-li pozorování uspořádaná podle velikosti $x_{(i)}$,

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

je testovou statistikou

$$D_n = \max \left\{ \left| F(x_{(i)}) - \frac{i-1}{n} \right|, \left| F(x_{(i)}) - \frac{i}{n} \right| \right\}, \quad i = 1, 2, \dots, n.$$

H_0 se zamítá, je-li $D_n \geq D_\alpha$

Kritické hodnoty D_α jsou tabelovány (Tab. 1)



Modifikovaný Kolmogorovův-Smirnovův test

Nejsou-li parametry normálního rozdělení známy (neúplně specifikovaný model), nahradí se odhady. Při rozhodování se musí použít jiné kritické hodnoty (Tab. 2).

Tab. 1 Kritické hodnoty $D_n(\alpha)$ maximální odchylky empirické distribuční funkce od teoretické

n	α					n	α				
	0.20	0.10	0.05	0.02	0.01		0.20	0.10	0.05	0.02	0.01
1	,90000	,95000	,97500	,99000	,99500	51	,14697	,16796	,18659	,20864	,22386
2	,68377	,77639	,84189	,90000	,92929	52	,14558	,16637	,18482	,20667	,22174
3	,56481	,63604	,70760	,78456	,82900	53	,14423	,16483	,18311	,20475	,21968
4	,49265	,56522	,62394	,68887	,73424	54	,14292	,16332	,18144	,20289	,21768
5	,44698	,50945	,56328	,62718	,66853	55	,14164	,16186	,17981	,20107	,21574
6	,41037	,46799	,51926	,57741	,61661	56	,14040	,16044	,17823	,19930	,21384
7	,38148	,43607	,48342	,53844	,57581	57	,13919	,15906	,17669	,19758	,21199
8	,35831	,40962	,45427	,50654	,54179	58	,13801	,15771	,17519	,19590	,21019
9	,33910	,38746	,43001	,47960	,51332	59	,13686	,15639	,17373	,19427	,20844
10	,32280	,36866	,40925	,45662	,48893	60	,13573	,15511	,17231	,19267	,20673
11	,30829	,35242	,39122	,43670	,46770	61	,13464	,15385	,17091	,19112	,20506
12	,29577	,33815	,37543	,41918	,44905	62	,13357	,15263	,16956	,18960	,20343
13	,28470	,32549	,36143	,40362	,43247	63	,13253	,15144	,16823	,18812	,20184
14	,27481	,31417	,34890	,38970	,41762	64	,13151	,15027	,16693	,18667	,20029
15	,26588	,30397	,33760	,37713	,40420	65	,13052	,14913	,16567	,18525	,19877
16	,25778	,29472	,32733	,36571	,39201	66	,12954	,14802	,16443	,18387	,19729
17	,25039	,28627	,31796	,35528	,38086	67	,12859	,14693	,16322	,18252	,19584
18	,24360	,27851	,30936	,34569	,37062	68	,12766	,14587	,16204	,18119	,19442
19	,23735	,27136	,30143	,33685	,36117	69	,12675	,14483	,16088	,17990	,19303

Tab. 2 Upravené kritické hodnoty dle Lilieforse

Rozsah výběru n	α				Rozsah výběru n	α			
	0,2	0,1	0,05	0,01		0,2	0,1	0,05	0,01
4	0,303	0,346	0,376	0,413	16	0,176	0,195	0,213	0,247
5	0,289	0,319	0,343	0,397	17	0,171	0,190	0,207	0,240
6	0,269	0,297	0,323	0,371	18	0,167	0,185	0,202	0,234
7	0,252	0,280	0,304	0,351	19	0,163	0,181	0,197	0,228
8	0,239	0,265	0,288	0,333	20	0,159	0,176	0,192	0,223
9	0,227	0,252	0,274	0,317	25	0,143	0,159	0,173	0,201
10	0,217	0,241	0,262	0,304	30	0,131	0,146	0,159	0,185
11	0,208	0,231	0,251	0,291	40	0,115	0,128	0,139	0,162
12	0,200	0,222	0,242	0,281	100	0,074	0,082	0,089	0,104
13	0,193	0,215	0,234	0,271	400	0,037	0,041	0,045	0,052
14	0,187	0,208	0,226	0,262	900	0,025	0,028	0,030	0,035
15	0,181	0,201	0,219	0,254					

PŘÍKLAD 2

Bylo provedeno $n = 12$ měření zatížení vlákna do přetržení:

i	x_i	i/n	$(i-1)/n$	$F(x)$	$ F(x)-(i-1)/n $	$ F(x)-i/n $
1	2,104	0,083333	0,000000	0,120882	0,120882	0,037548
2	2,222	0,166667	0,083333	0,200843	0,117509	0,034176
3	2,247	0,250000	0,166667	0,221160	0,054493	0,028840
4	2,286	0,333333	0,250000	0,255089	0,005089	0,078244
5	2,327	0,416667	0,333333	0,293492	0,039842	0,123175
6	2,367	0,500000	0,416667	0,333351	0,083315	0,166649
7	2,388	0,583333	0,500000	0,355096	0,144904	0,228237
8	2,512	0,666667	0,583333	0,490928	0,092405	0,175738
9	2,707	0,750000	0,666667	0,700509	0,033842	0,049491
10	2,751	0,833333	0,750000	0,742041	0,007959	0,091292
11	3,158	0,916667	0,833333	0,963648	0,130315	0,046982
12	3,172	1,000000	0,916667	0,966679	0,050012	0,033321

prumer	2,520083
rozptyl	0,126343
sm.odch	0,355447

max	0,130315	0,228237
-----	----------	----------

Závěr : Vzhledem k tomu, že maximální absolutní diference mezi empirickou distribuční funkcí a teoretickou distribuční funkcí není větší než kritická hodnota $D_n(\alpha)$, nemáme důvod zamítnout testovanou hypotézu H_0 : výběr pochází ze základního souboru s normálním rozdělením.



Grafický test

Do pravděpodobnostního papíru zakreslíme průběh empirické distribuční funkce, tj. body $[x_{(i)} ; i/n]$ a přímku odpovídající průběhu odhadu distribuční funkce rozdělení $N(\mu, \sigma^2)$ $\hat{F}(x)$.

K odhadu teoretické distribuční funkce zakreslíme meze konfidenčního intervalu, tj. body

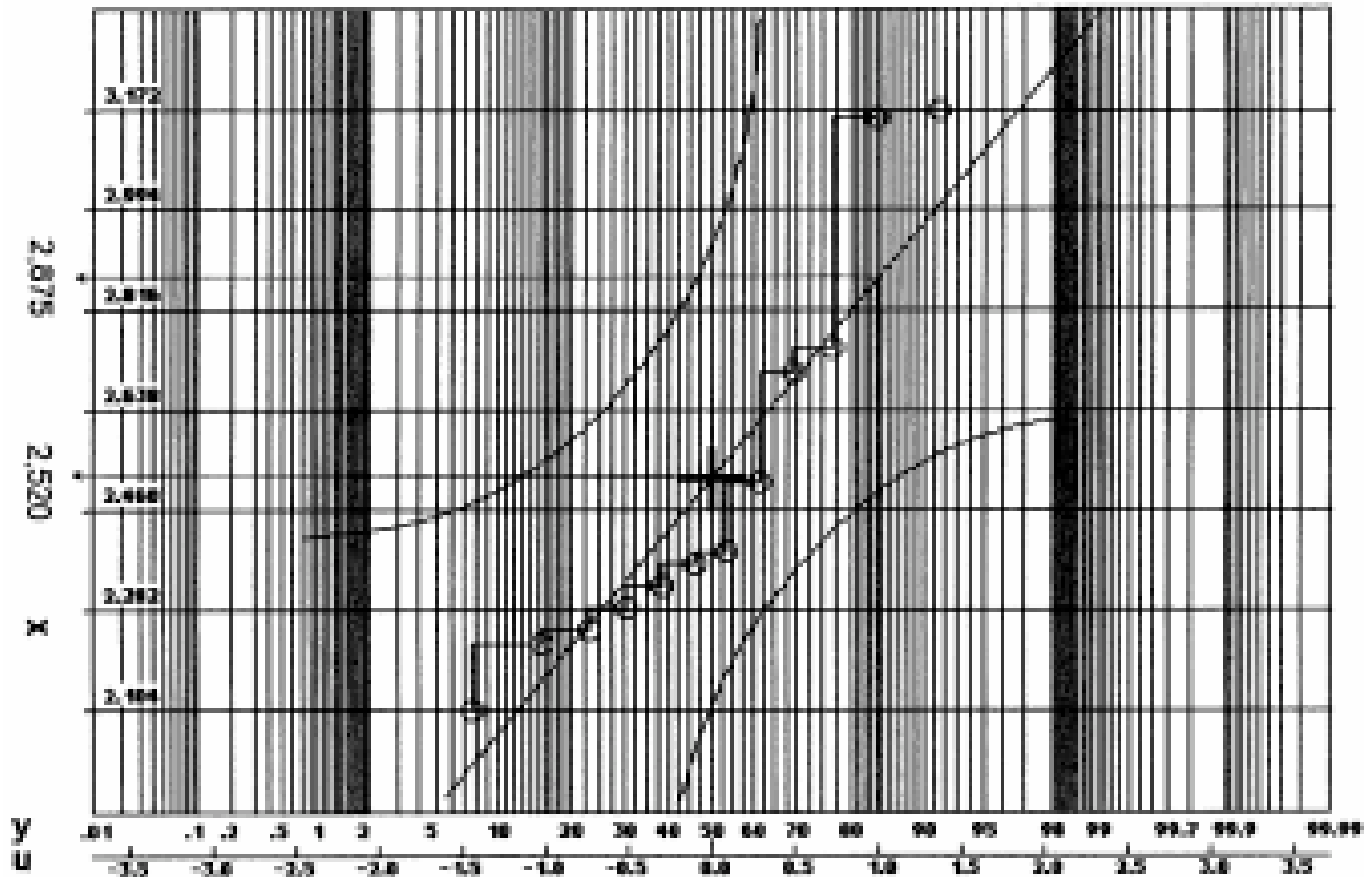
$$[x ; \hat{F}(x) \pm D_n(\alpha)] .$$

Vzniklé dvě křivky představují konfidenční interval distribuční funkce $F(x)$ s konfidenční úrovní $1-\alpha$.

Testovaná hypotéza H_0 se zamítá, na hladině významnosti α , jestliže alespoň pro jednu hodnotu x empirická distribuční funkce, znázorněná na grafu body, leží vně zakresleného pásma.

Aplikace testu normality, pomocí pravděpodobnostního papíru

$\alpha = 0.05$





PŘÍKLAD 2 pokračování:

Na obrázku je vedle pravděpodobnostní stupnice $y = 100 F(x)$ ještě stupnice u , odpovídající kvantilům normovaného normálního rozdělení $N(0, 1)$. (Platí tedy $100 \phi(u) = y$.)

Přímku představující odhad distribuční funkce hypotetického normálního rozdělení $N(\mu = 2,520 ; \sigma^2 = 0,355^2)$ proložíme body

$$(\bar{x} = 2,520 ; u = 0) \text{ a } (\bar{x} + s = 2,875 ; u = 1) .$$

Pro $n = 12$ a $\alpha = 0,05$ je $D_n(\alpha) = D_{12}(0,05) = 0,37543$. Tedy hranice zakreslené na obrázku jsou $(F(x) \pm 0,375) * 100$.

Závěr : Ani jeden bod neleží mimo zakreslené meze, nemáme důvod zamítnout testovanou hypotézu H_0 .



Testy normality v MINITABu

- n Kolmogorov – Smirnov
- n Anderson – Darling
 - n testová statistika A^2 (A squared)
hodnoty větší než kritické svědčí proti normalitě
- n Ryan – Joiner
 - n testová statistika R
podobný Shapiro-Wilkově testu (viz dále)
hodnoty menší než kritické svědčí proti normalitě



Použití p-hodnoty

n Na výstupu každé procedury pro statistický test je kromě hodnoty testové statistiky uvedena tzv. **p-hodnota** (p-value)

n Platí-li:

$$p\text{-hodnota} < \alpha,$$

zamítneme testovanou hypotézu na hladině významnosti α .



Pravděpodobnostní graf v MINITABu

osa x – naměřené hodnoty $x_{(i)}$ sledované veličiny uspořádané podle velikosti

osa y – hodnoty empirické distribuční funkce vynášené na nelineární stupnici, vycházející z předpokladu normality

y-ová souřadnice bodu odpovídá kvantilu $u_{(i)}$ rozdělení $N(0,1)$

červeně proložena regresní přímka

$$E\{x_{(i)}\} = m + S u_{(i)}$$

Normálnímu rozdělení veličiny X odpovídají vynesené body ležící v blízkosti přímky a nevykazující nápadný nelineární trend.

Graf je buď doplněn výsledkem některého z uvedených testů normality nebo 95% pásem spolehlivosti.



Testy normality ve Statistice

- n chí-kvadrát
- n Kolmogorov – Smirnov
- n Shapiro-Wilk
 - n testová statistika W
čím blíže 1, tím více svědčí pro normalitu



Grafické metody ve Statistice

- n pravděpodobnostní graf
 - n osa x – naměřené hodnoty $x_{(i)}$ seřazené podle velikosti
 - n osa y – kvantily $u_{(i)}$ rozdělení $N(0,1)$



Q - Q graf

- n osa x - kvantily $u_{(i)}$ rozdělení $N(0,1)$
- n osa y - naměřené hodnoty $x_{(i)}$ seřazené podle velikosti
- n vyneseny body je proložena regresní přímka
- n z rovnice regresní přímky se odhadnou parametry



P - P graf

- n osa x – hodnoty teoretické distribuční funkce (lineární stupnice)
- n osa y – hodnoty empirické distribuční funkce (lineární stupnice)
- n v grafu vyznačena přímka se směrnici 1



Výhoda grafických metod

- n Naznačují, o jaké rozdělení se ve skutečnosti jedná. I v případě, že testy vycházejí nevýznamné, může nelineární trend v grafu prozradit vhodnost jiného než normálního rozdělení.
- n Někdy umožňují lépe posoudit, zda nepřijatelnost hypotézy o normalitě je důsledkem existence několika extrémních pozorování, nebo zda je výběrové rozdělení skutečně jiné než normální.



PŘÍKLAD 3

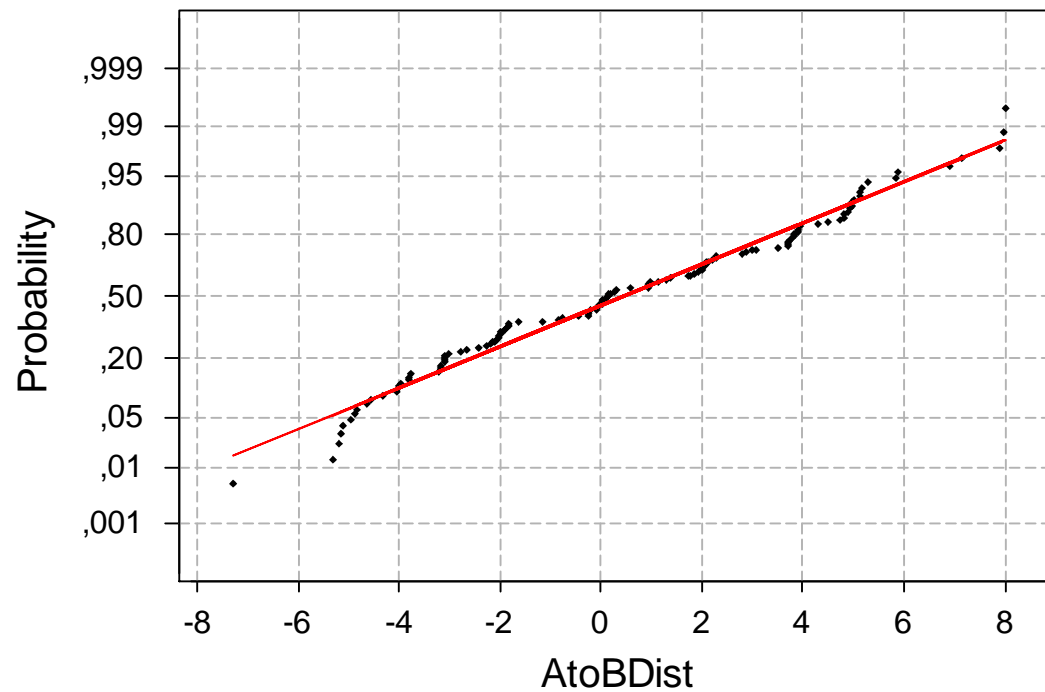
V rámci SPC se v montážním závodě kontroluje vzdálenost aktuální pozice bodu na klikovém hřídeli od základní pozice.

Každý den se provedlo 5 měření, k dispozici jsou hodnoty za 25 dní.

Před výpočtem indexu způsobilosti je třeba ověřit, zda lze rozdělení hodnot měřené vzdálenosti považovat za normální.

Příklad 3 - MINITAB

Normal Probability Plot



Average: 0,441704
StDev: 3,49136
N: 125

Anderson-Darling Normality Test
A-Squared: 0,891
P-Value: 0,022



Příklad 3 - Výsledky různých testů normality

Anderson-Darling Normality Test

A-Squared: 0,891

P-Value: 0,022

Ryan-Joiner Normality Test

R: 0,9900

P-Value (approx): 0,0656

Kolmogorov-Smirnov Normality Test

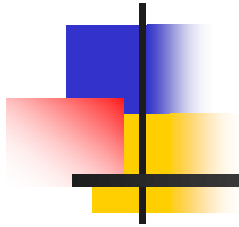
D+: 0,094 D-: 0,074 D : 0,094

Approximate P-Value < 0.01

Shapiro-Wilk Normality Test

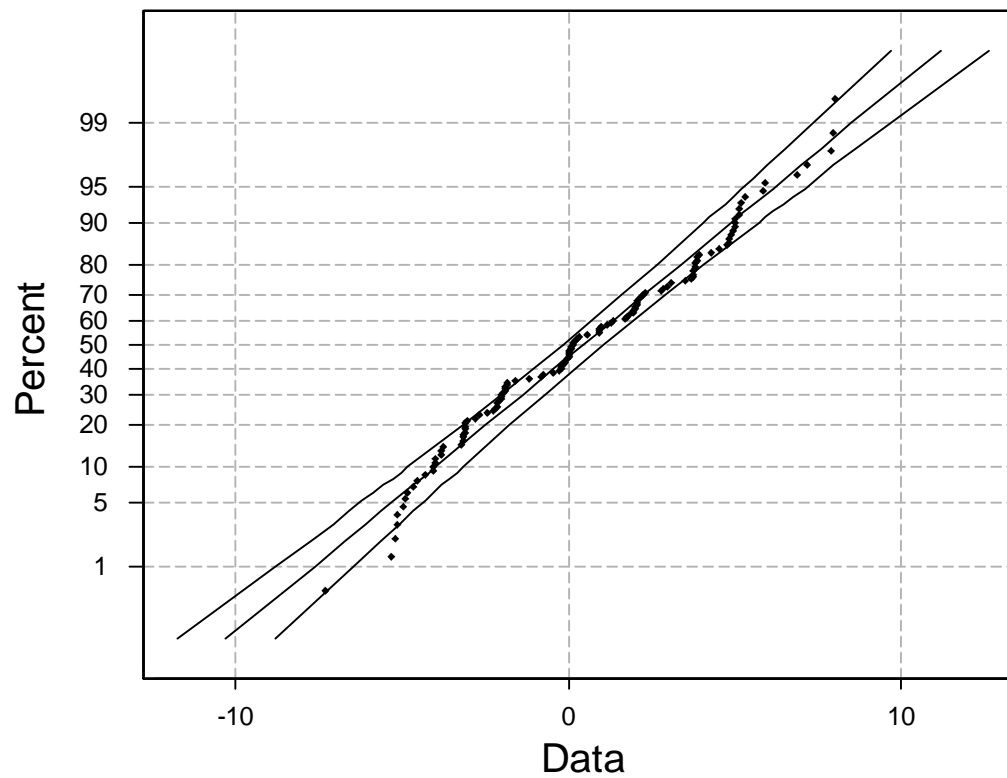
W: 0,976

P-Value: 0,0279



Příklad 3 - MINITAB

Normal Probability Plot for AtoBDist
ML Estimates - 95% CI



ML Estimates

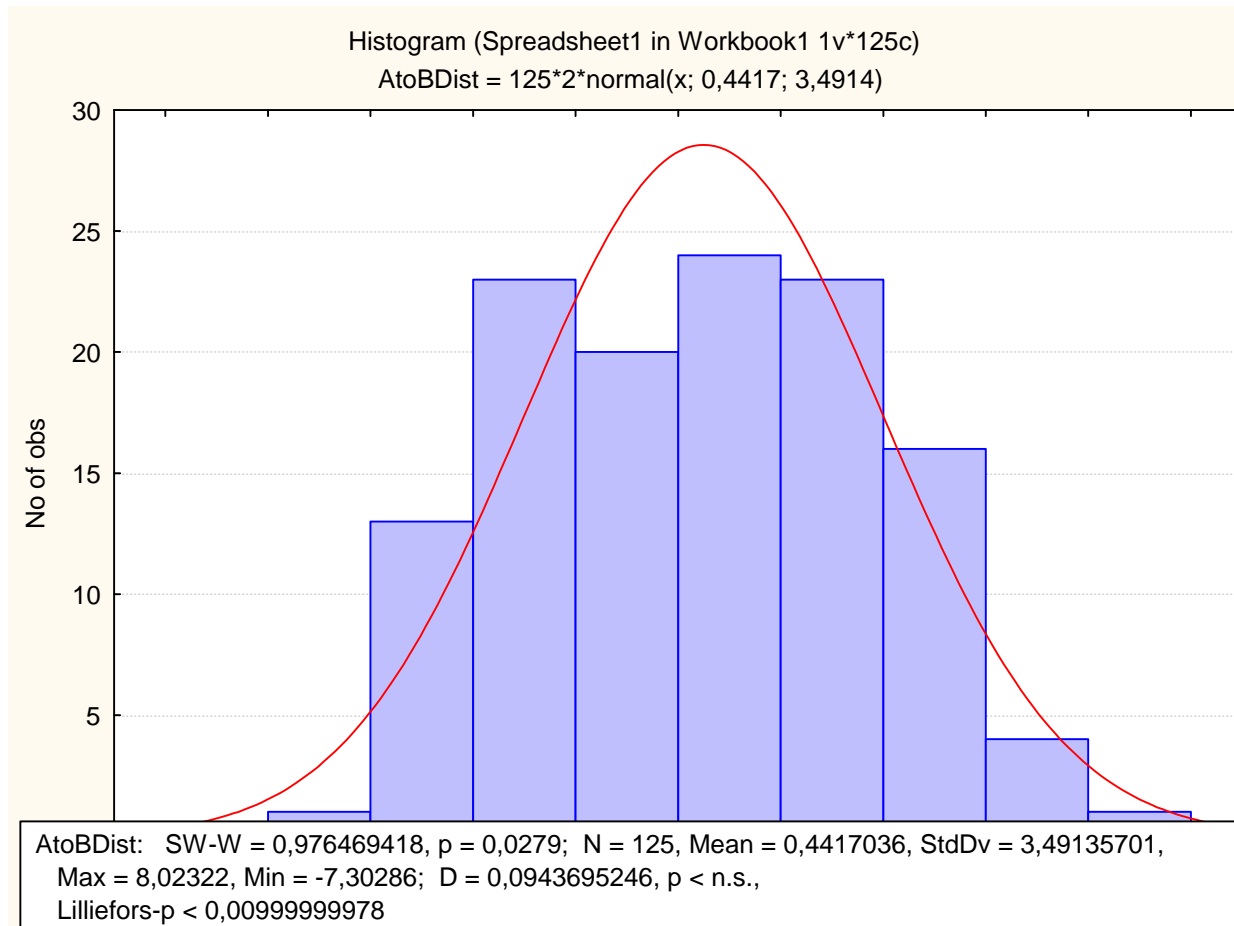
Mean 0,441704

StDev 3,47736

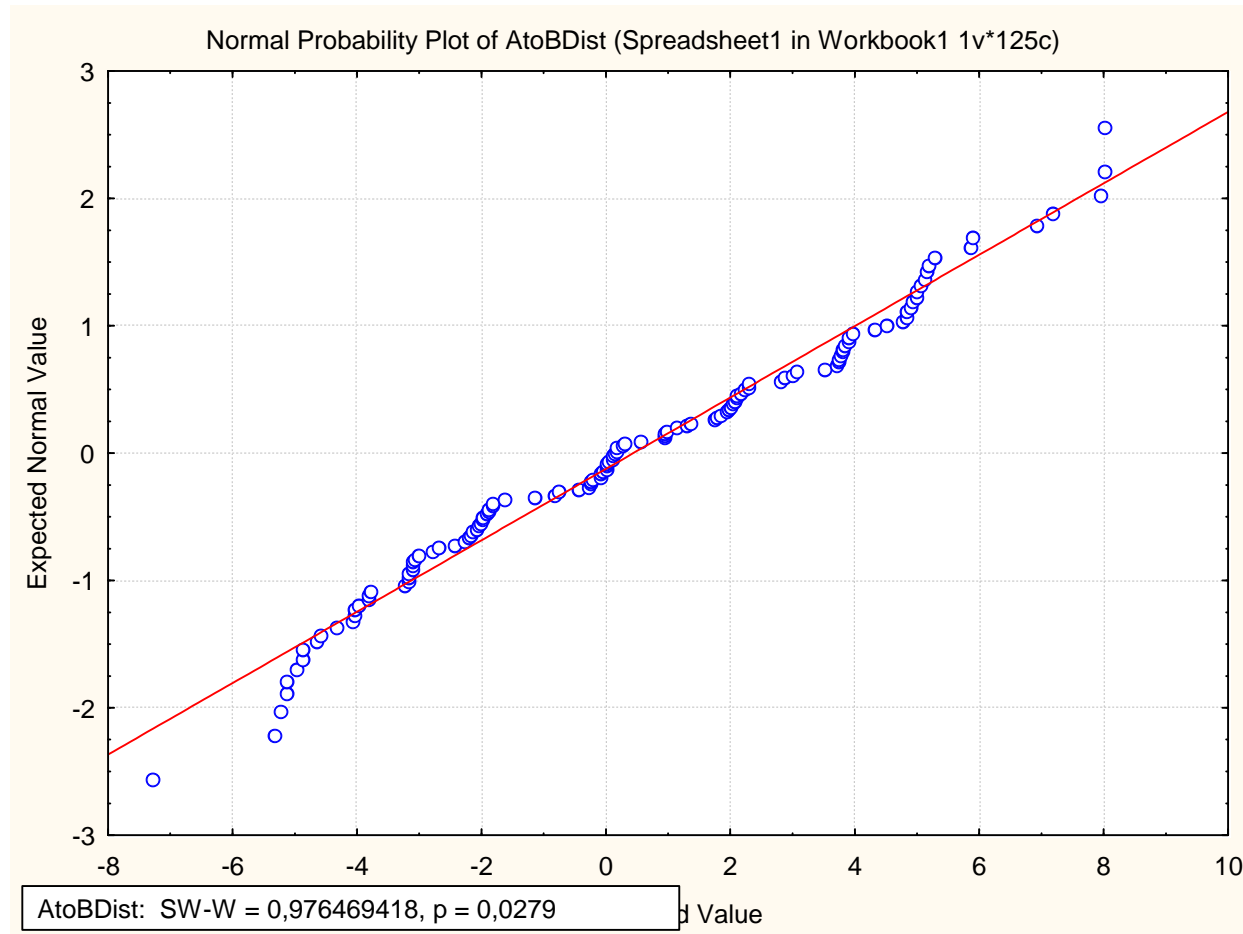
Goodness of Fit

AD* 1

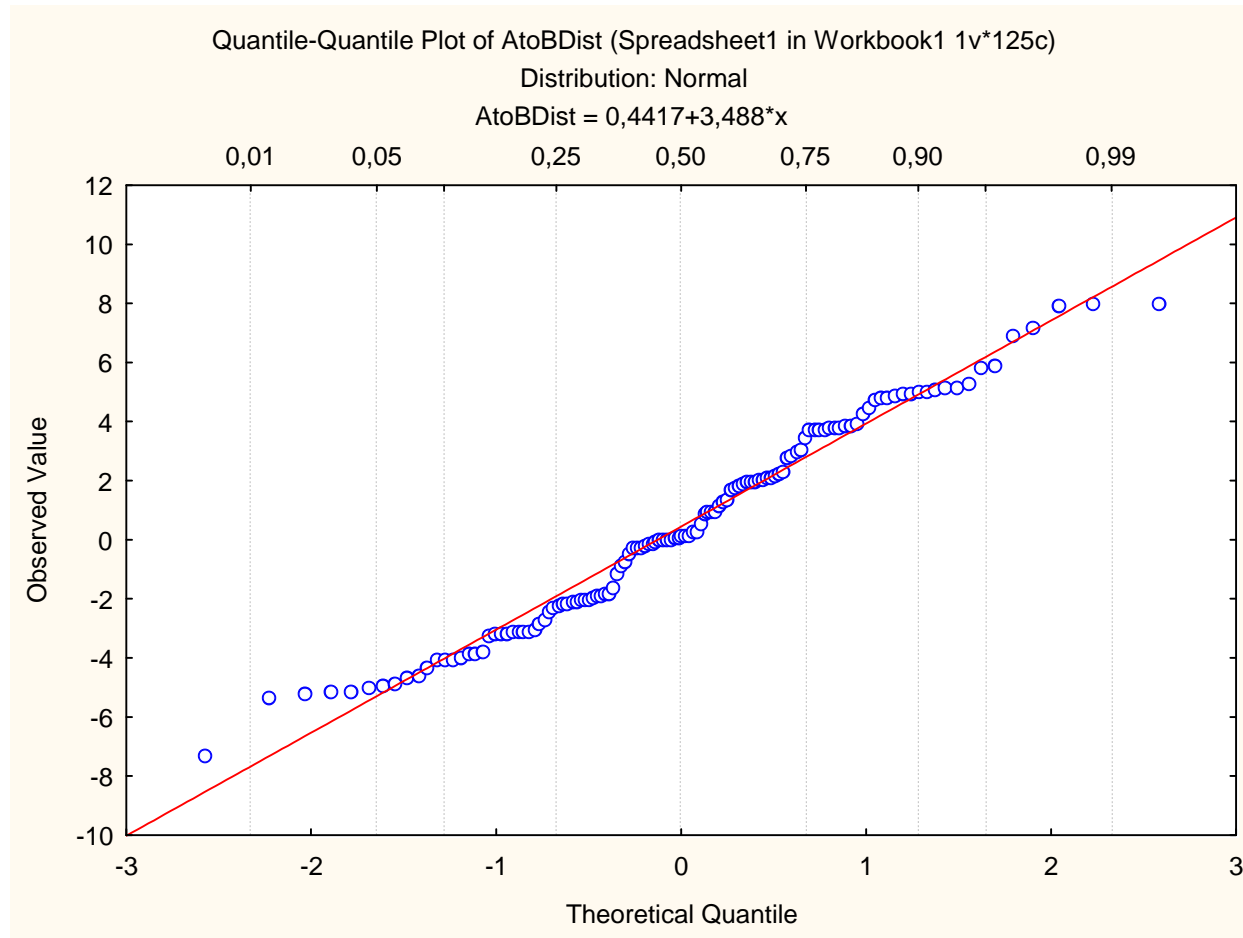
Příklad 3 - Statistica



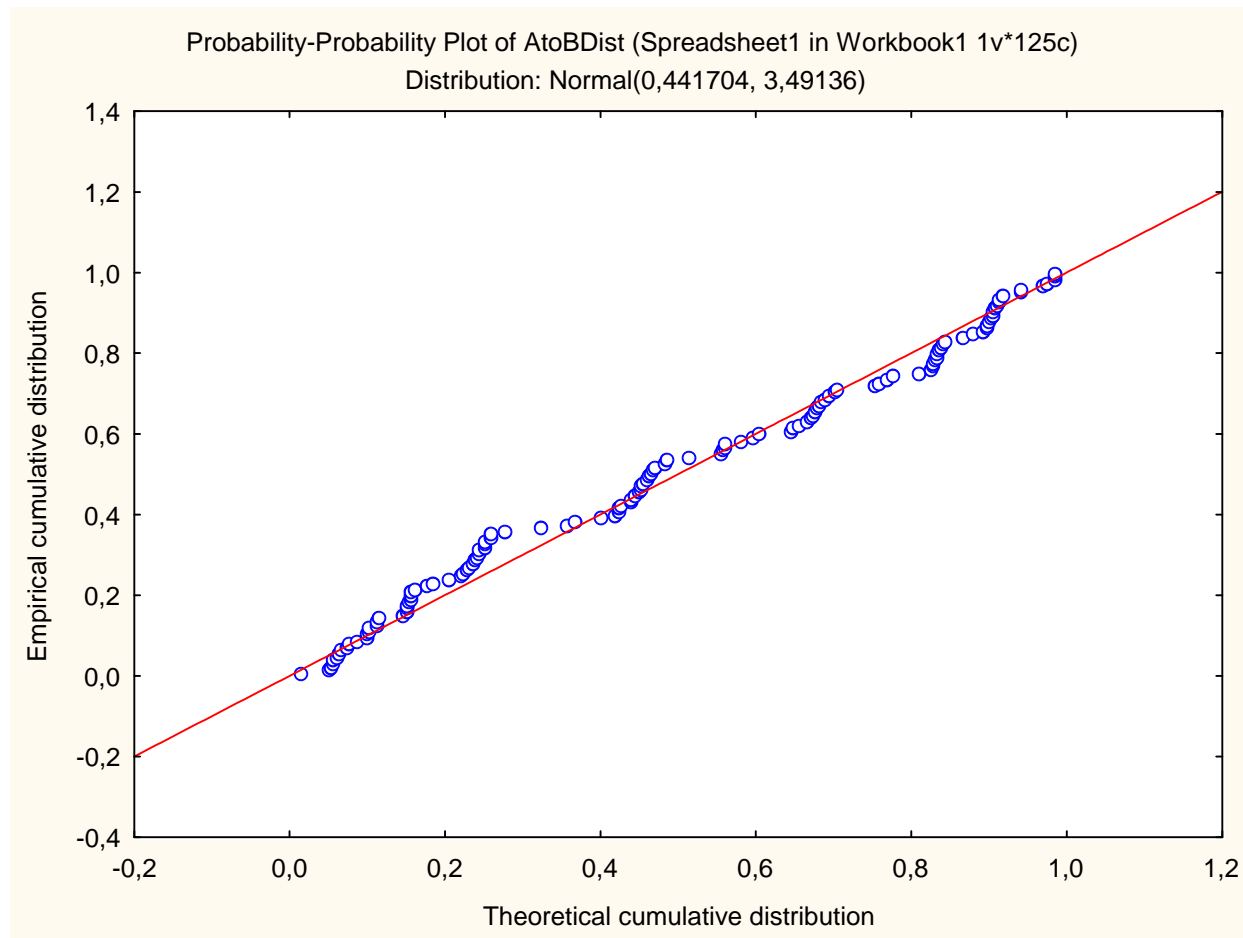
Příklad 3 - Statistica



Příklad 3 - Statistica



Příklad 3 - Statistica



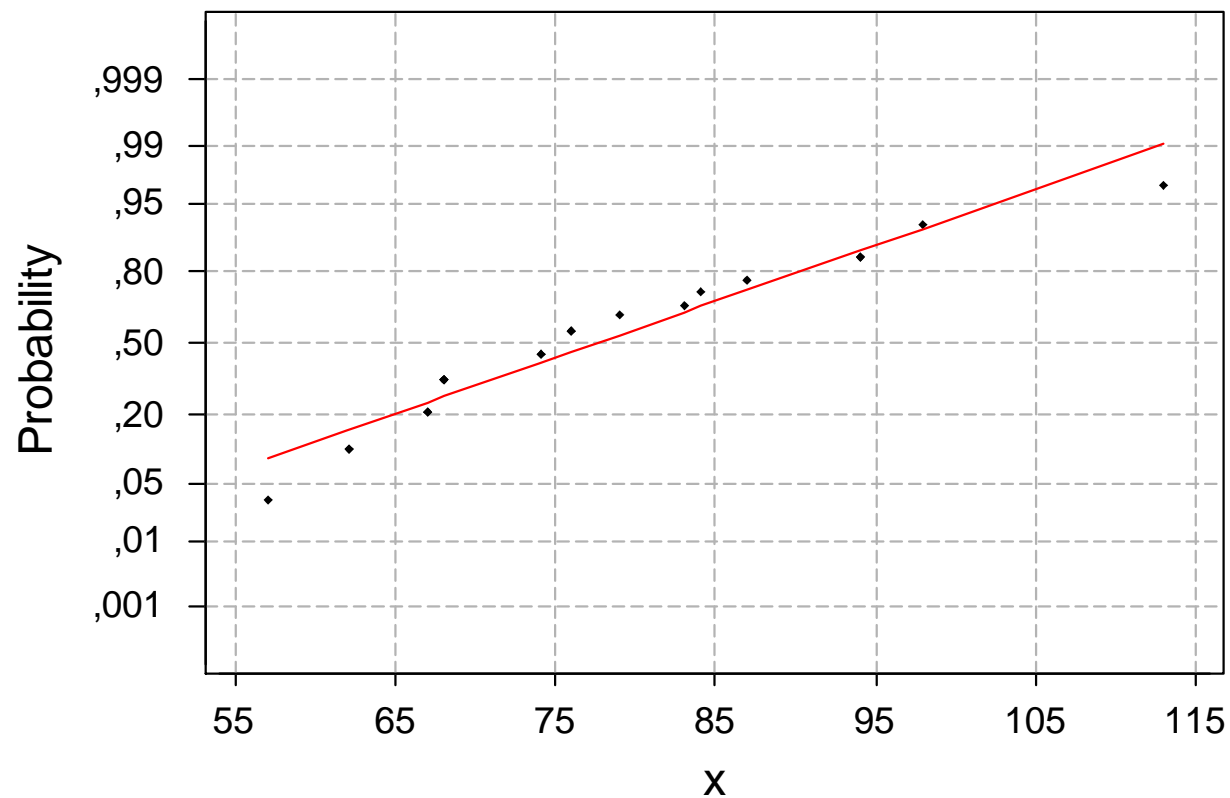


Příklad 4

- n Při elektronickém testování ozubených kol se sleduje maximální odchylka profilu od ideálního tvaru.

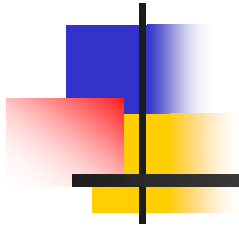
Příklad 4 - MINITAB

Normal Probability Plot



Average: 77,55
StDev: 14,1625
N: 20

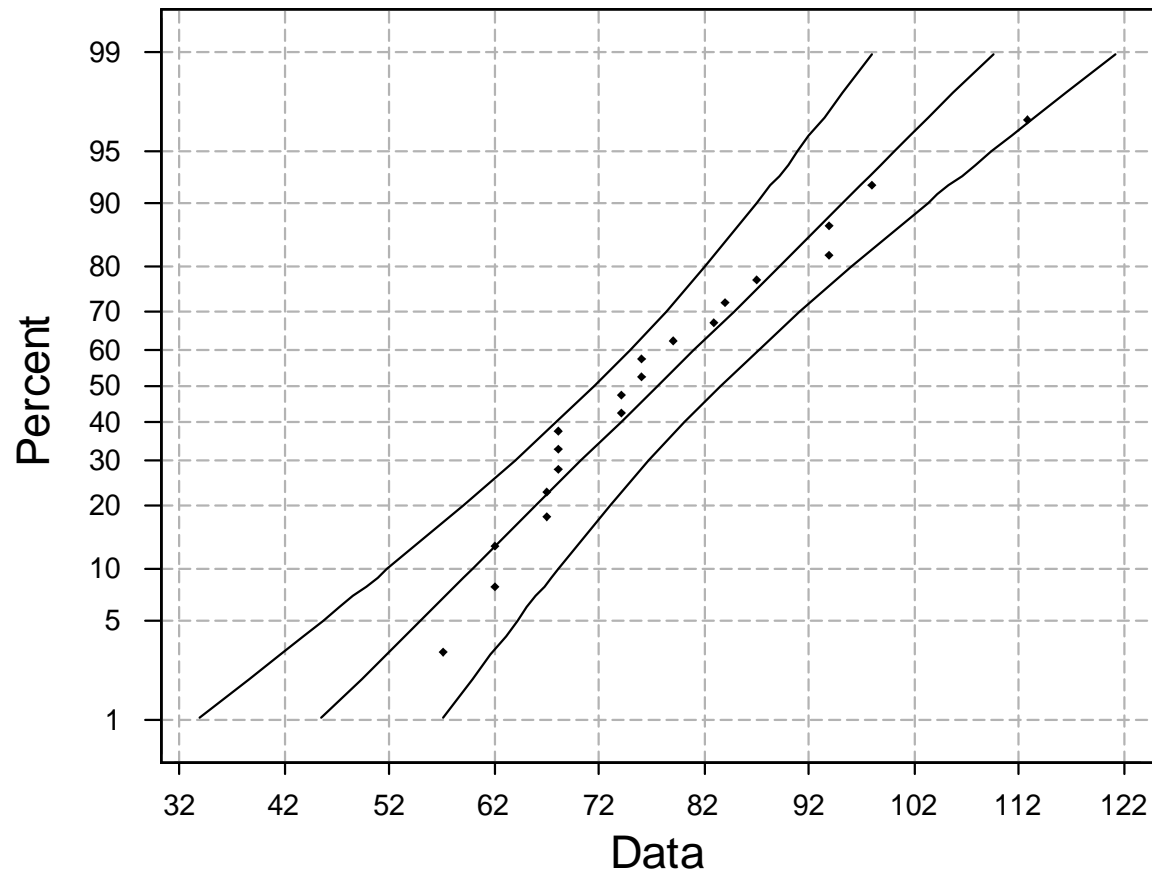
Anderson-Darling Normality Test
A-Squared: 0,455
P-Value: 0,240



Příklad 4

Normal Probability Plot for x

ML Estimates - 95% CI



ML Estimates

Mean 77,55

StDev 13,8039

Goodness of Fit

AD* 0,997



Příklad 4 - Výsledky různých testů normality

Anderson-Darling Normality Test

A-Squared: 0,455

P-Value: 0,240

Ryan-Joiner Normality Test

R: 0,9721

P-Value (approx): $\geq 0,1000$

Kolmogorov-Smirnov Normality Test

D+: 0,119 D-: 0,073 D : 0,119

Approximate P-Value $\geq 0,15$

Shapiro-Wilk Normality Test

W: 0,939

P-Value: 0,2313

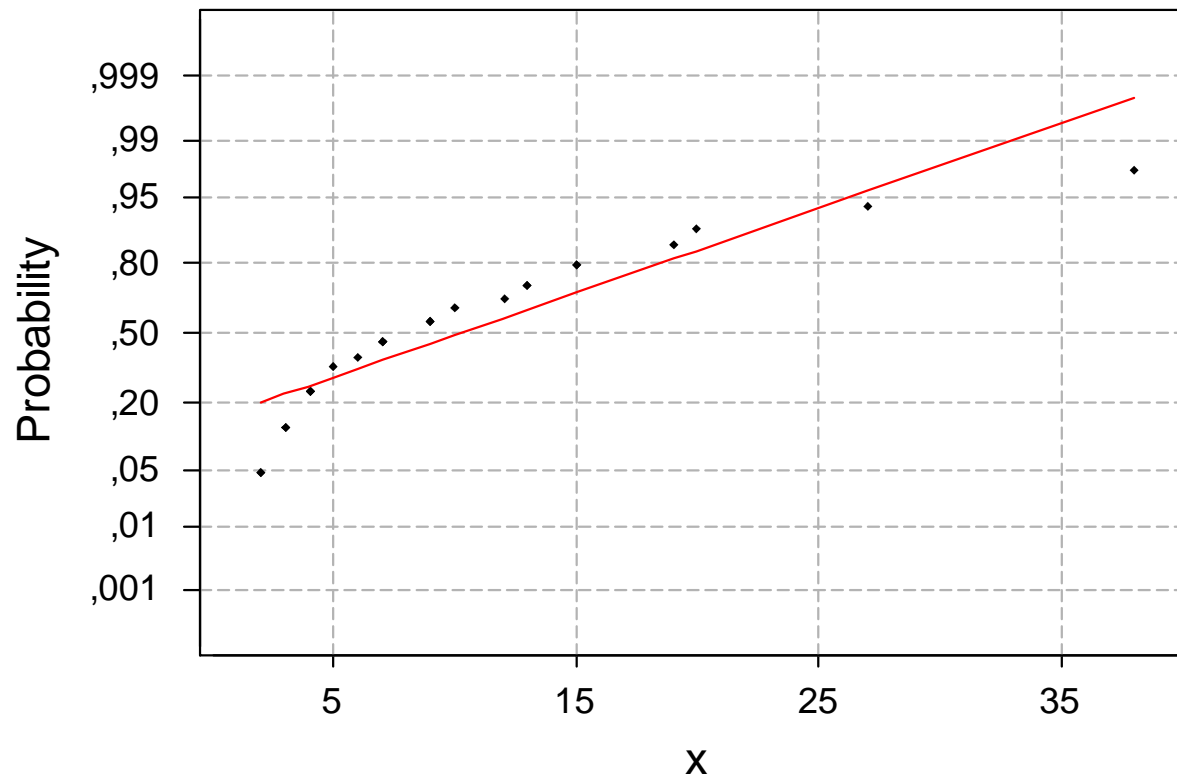


Příklad 5

- n 25 vzorků materiálu pro operační přístroje bylo testováno na obsah kovových příměsí.

Příklad 5 - MINITAB

Normal Probability Plot



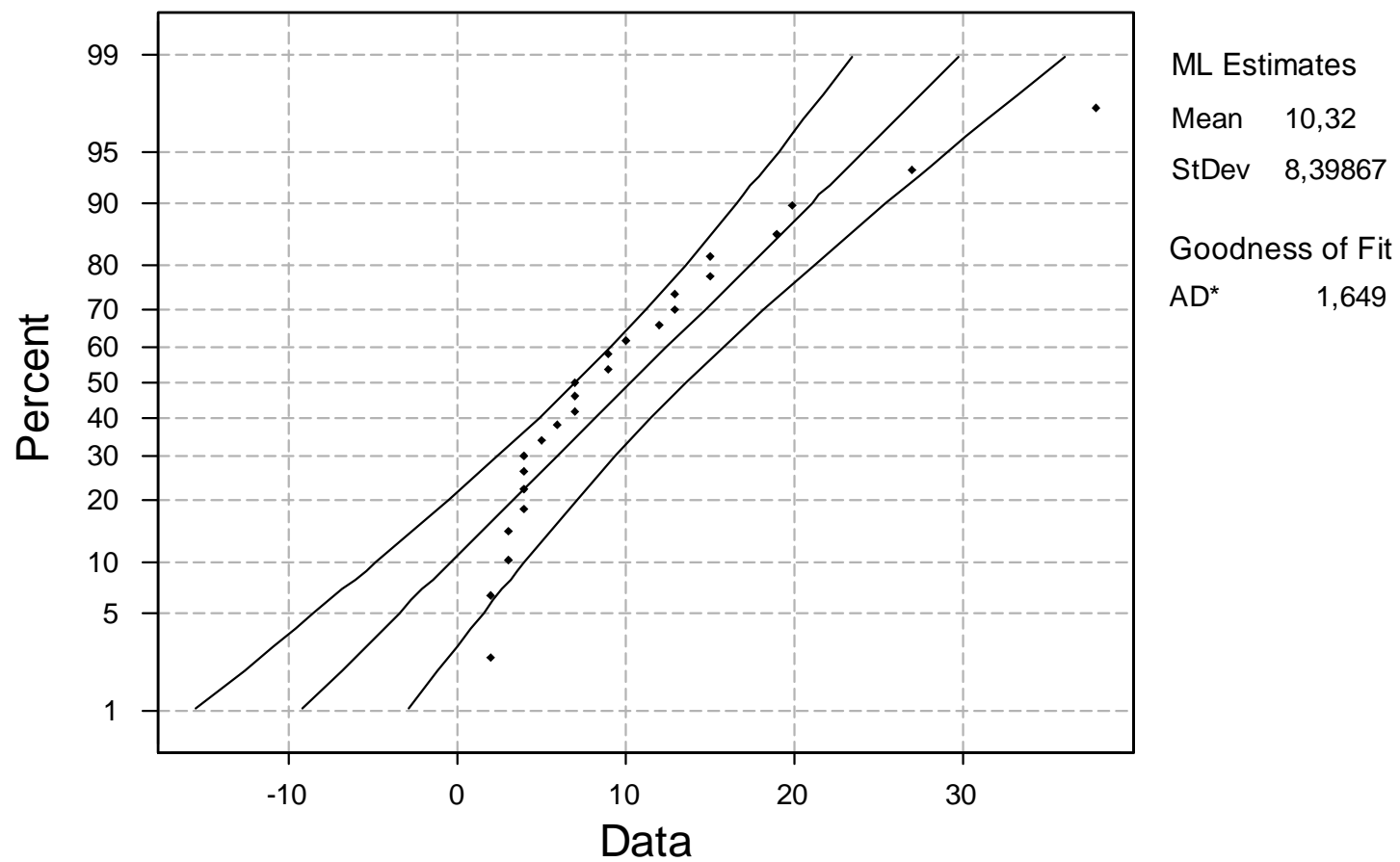
Average: 10,32
StDev: 8,57185
N: 25

Anderson-Darling Normality Test
A-Squared: 1,276
P-Value: 0,002

Příklad 5 - MINITAB

Normal Probability Plot for x

ML Estimates - 95% CI





Příklad 5 - Výsledky různých testů normality

Anderson-Darling Normality Test

A-Squared: 1,276

P-Value: 0,002

Ryan-Joiner Normality Test

R: 0,9131

P-Value (approx): $< 0,0100$

Kolmogorov-Smirnov Normality Test

D+ : 0,155 D- : 0,146 D : 0,155

Approximate P-Value: 0,122

Shapiro-Wilk Normality Test

W: 0,827

P-Value: 0,0007



Příloha – vzorce 1

- Anderson - Darling

$$A^2 = -\frac{1}{n} \sum_{i=1}^n (2i-1) [\ln \Phi_i + \ln(1-\Phi_{n-i+1})] - n$$

$$\Phi_i = \Phi(u_{(i)}) \quad u_{(i)} = \frac{x_{(i)} - \bar{x}}{\hat{S}}$$

$$\hat{S}^2 = \frac{1}{n} \sum_{i=1}^n (x_{(i)} - \bar{x})^2$$

maximálně věrohodný odhad σ^2



Příloha – vzorce 2

- Shapiro - Wilk

$$W = \frac{\left(\sum u_{(i)} x_{(i)}\right)^2}{\sum u_{(i)}^2 \sum (x_{(i)} - \bar{x})^2}$$

$$u_{(i)} = \Phi^{-1}\left(\frac{i - 3/8}{n + 1/4}\right)$$