

MV011 Statistika I – cvičení 13

1) Testujte hypotézu (v SAS pomocí Proc Univariate a Proc Ttest), že průměrné SAT score je rovno 1200. Data získáte ze sloupce SATScore tabulky testscores vytvořené pomocí gen_data_reg.sas. Následně otestujte hypotézu, že průměrné score je rovno 1000.

a) v SASu

a1) $\mu_0 = 1200$

Testing Whether the Mean of SAT Scores = 1200

The UNIVARIATE Procedure
Variable: SATScore

Tests for Location: Mu0=1200				
Test	Statistic		p Value	
Student's t	t	-0.5702	Pr > t	0.5702
Sign	M	-5	Pr >= M	0.3019
Signed Rank	S	-207	Pr >= S	0.2866

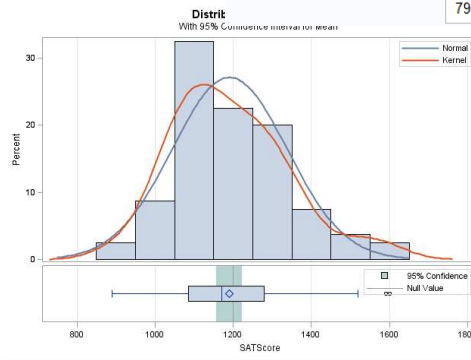
Testing Whether the Mean of SAT Scores = 1200 Using PROC TTEST

The TTEST Procedure
Variable: SATScore

N	Mean	Std Dev	Std Err	Minimum	Maximum
80	1190.6	147.1	16.4416	890.0	1600.0

Mean	95% CL Mean	Std Dev	95% CL Std Dev
1190.6	1157.9 1223.4	147.1	127.3 174.2

DF	t Value	Pr > t
79	-0.57	0.5702



a1) $\mu_0 = 1100$

Testing Whether the Mean of SAT Scores = 1100

The UNIVARIATE Procedure
Variable: SATScore

Tests for Location: Mu0=1100				
Test	Statistic		p Value	
Student's t	t	5.511922	Pr > t	<.0001
Sign	M	16.5	Pr >= M	0.0002
Signed Rank	S	931.5	Pr >= S	<.0001

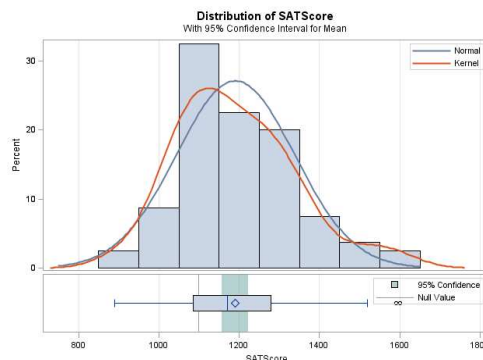
Testing Whether the Mean of SAT Scores = 1100 Using PROC TTEST

The TTEST Procedure
Variable: SATScore

N	Mean	Std Dev	Std Err	Minimum	Maximum
80	1190.6	147.1	16.4416	890.0	1600.0

Mean	95% CL Mean	Std Dev	95% CL Std Dev
1190.6	1157.9 1223.4	147.1	127.3 174.2

DF	t Value	Pr > t
79	5.51	<.0001



2) Testujte hypotézu (v SAS pomocí Proc Ttest), že průměrné SAT score chlapců a dívek je stejný. Data získáte ze sloupce SATScore tabulky testscores vytvořené pomocí gen_data_reg.sas. Chlapci a dívky jsou v datech určení sloupcem Gender.

Two-Sample t-test Comparing Girls to Boys

The TTEST Procedure

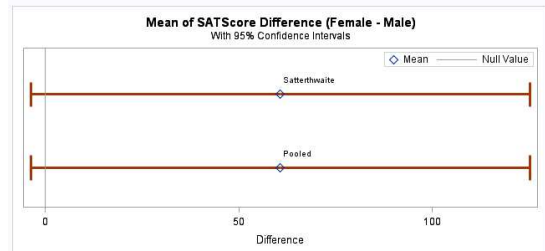
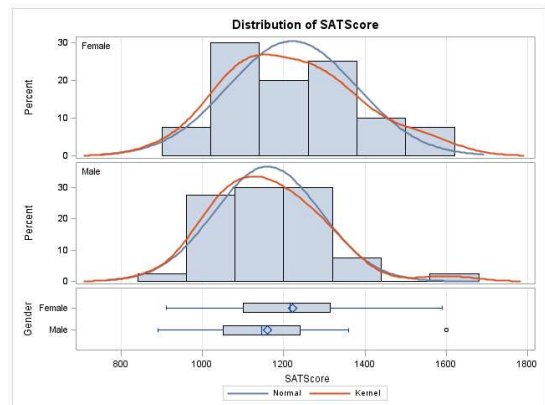
Variable: SATScore

Gender	N	Mean	Std Dev	Std Err	Minimum	Maximum
Female	40	1221.0	157.4	24.8864	910.0	1590.0
Male	40	1160.3	130.9	20.7008	890.0	1600.0
Diff (1-2)		60.7500	144.8	32.3706		

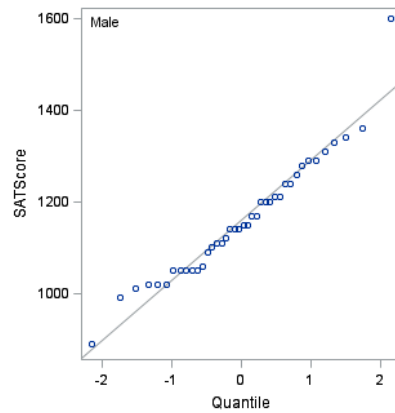
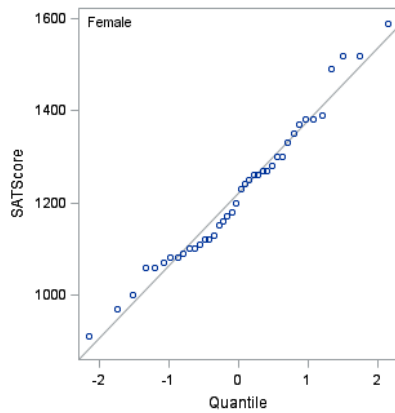
Gender	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
Female		1221.0	1170.7 1271.3	157.4	128.9 202.1
Male		1160.3	1118.4 1202.1	130.9	107.2 168.1
Diff (1-2)	Pooled	60.7500	-3.6950 125.2	144.8	125.2 171.7
Diff (1-2)	Satterthwaite	60.7500	-3.7286 125.2		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	78	1.88	0.0643
Satterthwaite	Unequal	75.497	1.88	0.0644

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	39	39	1.45	0.2545



Q-Q Plots of SATScore



3) Testujte hypotézu o nezávislosti pohlaví a přežití na Titaniku. Dále pak testujte hypotézu o nezávislosti palubní třídy a přežití na Titaniku. Data získáte st100d01.sas – tabulka Titanic, sloupce Gender, Class a Survived.

Associations with Survival				
The FREQ Procedure				
Frequency Percent Row Pct Col Pct	Table of Gender by Survived			
	Gender	Survived		
		Died	Survived	Total
	female	127	339	466
		9.70	25.90	35.60
		27.25	72.75	
		15.70	67.80	
	male	682	161	843
		52.10	12.30	64.40
		80.90	19.10	
84.30		32.20		
Total	809	500	1309	
	61.80	38.20	100.00	
Statistics for Table of Gender by Survived				
Statistic	DF	Value	Prob	
Chi-Square	1	365.8869	<.0001	
Likelihood Ratio Chi-Square	1	372.9213	<.0001	
Continuity Adj. Chi-Square	1	363.6179	<.0001	
Mantel-Haenszel Chi-Square	1	365.6074	<.0001	
Phi Coefficient		-0.5287		
Contingency Coefficient		0.4674		
Cramer's V		-0.5287		
Frequency Percent Row Pct Col Pct	Table of Class by Survived			
	Class	Survived		
		Died	Survived	Total
	1	123	200	323
		9.40	15.28	24.68
		38.08	61.92	
		15.20	40.00	
	2	158	119	277
		12.07	9.09	21.16
		57.04	42.96	
19.53		23.80		
3	528	181	709	
	40.34	13.83	54.16	
	74.47	25.53		
	65.27	36.20		
Total	809	500	1309	
	61.80	38.20	100.00	
Statistics for Table of Class by Survived				
Statistic	DF	Value	Prob	
Chi-Square	2	127.8592	<.0001	
Likelihood Ratio Chi-Square	2	127.7655	<.0001	
Mantel-Haenszel Chi-Square	1	127.7093	<.0001	
Phi Coefficient		0.3125		
Contingency Coefficient		0.2983		
Cramer's V		0.3125		

4) Viz přednáškový text str. 659