Statistics for Analytics (BAN 100)

Assignment 2

TECHNICAL DESK

by : Aaron Gonsalves

(161288196)

HYPOTHESIS 1

- Compare average weight of babies of black mother and non-black mother.

Null Hypothesis(H_0):	**Average weight of babies of black mother is equal to average weight of babies of non-black mother.**
Alternative Hypothesis(H_a):	**Average weight of babies of black mother is not equal to average weight of babies of non-black mother.**

T-TEST RESULTS

```
title 'T-test on weight by class : black (mother is black or non-black)';
proc ttest data=work.birth;
class Black;
var Weight;
run;
```

T-test on weight by class : black (mother is black or non-black)

The TTEST Procedure

Variable: Weight (Weight)

Black	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		41858	3411.2	547.6	2.6766	284.0	5970.0
1		8142	3162.7	613.7	6.8011	240.0	6350.0
Diff (1-2)	Pooled		248.6	558.9	6.7697		
Diff (1-2)	Satterthwaite		248.6		7.3088		

Black	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3411.2	3406.0	3416.5	547.6	543.9	551.4
1		3162.7	3149.3	3176.0	613.7	604.4	623.3
Diff (1-2)	Pooled	248.6	235.3	261.8	558.9	555.5	562.4
Diff (1-2)	Satterthwaite	248.6	234.2	262.9			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	49998	36.72	<.0001
Satterthwaite	Unequal	10808	34.01	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	8141	41857	1.26	<.0001

Interpretation

- In the 4th table of T-test results, the P value for F test is <.0001. This means that the P value is <0.05, So Null hypothesis will be rejected, implying that the variances are unequal.

- In the 3rd table of T-test results, we check null hypothesis of unequal means with assumption of unequal variances. Here, we find that the t-statistic for unequal variances is 34.01 with a degree of freedom 10808, along with P-value<.0001, inferring that it is not significant.

- As P-value<0.05, Null hypothesis (H_0) is rejected and there is much evidence in support of Alternative Hypothesis(H_a).

Conclusion

- As Null hypothesis (H_0) is rejected from our interpretation, the conclusion is in support of Alternative Hypothesis(H_a).

- Hence, it can be deduced that average weight of babies of black mother is not equal to average weight of babies of non-black mother.

- We also find that the average weight of babies of black mother is less than average weight of babies of non-black mother.

HYPOTHESIS 2

- Compare average weight of babies of married mother and unmarried mother.

Null Hypothesis(H_0):	**Average weight of babies of married mother is equal to average weight of babies of unmarried mother.**
Alternative Hypothesis(H_a):	**Average weight of babies of married mother is not equal to average weight of babies of unmarried mother.**

T-TEST RESULTS

```
title 'T-test on weight by class : married (mother is married or unmarried)';

proc ttest data=work.birth;

class Married;

var Weight;

run;
```

T-test on weight by class : married (mother is married or unmarried)

The TTEST Procedure

Variable: Weight (Weight)

Married	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		14369	3234.4	579.0	4.8302	284.0	6350.0
1		35631	3425.7	551.8	2.9231	240.0	5970.0
Diff (1-2)	Pooled		-191.3	559.7	5.5315		
Diff (1-2)	Satterthwaite		-191.3		5.6459		

Married	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3234.4	3225.0	3243.9	579.0	572.4	585.8
1		3425.7	3420.0	3431.5	551.8	547.8	555.9
Diff (1-2)	Pooled	-191.3	-202.1	-180.5	559.7	556.3	563.2
Diff (1-2)	Satterthwaite	-191.3	-202.4	-180.2			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	49998	-34.58	<.0001
Satterthwaite	Unequal	25443	-33.88	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	14368	35630	1.10	<.0001

Interpretation

- In the 4th table of T-test results, the P value for F test is <.0001. This means that the P value is <0.05, So Null hypothesis will be rejected, implying that the variances are unequal.

- In the 3rd table of T-test results, we check null hypothesis of unequal means with assumption of unequal variances. Here, we find that the t-statistic for unequal variances is -33.88 with a degree of freedom 25443, along with P-value<.0001, inferring that it is not significant.

- As P-value<0.05, Null hypothesis (H_0) is rejected and there is much evidence in support of Alternative Hypothesis(H_a).

Conclusion

- As Null hypothesis (H_0) is rejected from our interpretation, the conclusion is in support of Alternative Hypothesis(H_a).

- Hence, it can be deduced that average weight of babies of married mother is not equal to average weight of babies of unmarried mother.

- We also find that average weight of babies of married mother is greater than that of unmarried mother.

HYPOTHESIS 3

- Compare average weight of baby boys with baby girls.

Null Hypothesis(H_0):	**Average weight of baby boys is equal to average weight of baby girls.**
Alternative Hypothesis(H_a):	**Average weight of baby boys is not equal to average weight of baby girls.**

T-TEST RESULTS

```
title 'T-test on weight by class : boy (child is boy or girl)';
proc ttest data=work.birth;
class Boy;
var Weight;
run;
```

T-test on weight by class : boy (child is boy or girl)

The TTEST Procedure

Variable: Weight (Weight)

Boy	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		24208	3310.6	547.7	3.5204	240.0	6350.0
1		25792	3427.3	577.7	3.5970	284.0	5970.0
Diff (1-2)	Pooled		-116.7	563.4	5.0416		
Diff (1-2)	Satterthwaite		-116.7		5.0331		

Boy	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3310.6	3303.7	3317.5	547.7	542.9	552.7
1		3427.3	3420.2	3434.3	577.7	572.7	582.7
Diff (1-2)	Pooled	-116.7	-126.6	-106.8	563.4	559.9	566.9
Diff (1-2)	Satterthwaite	-116.7	-126.6	-106.8			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	49998	-23.15	<.0001
Satterthwaite	Unequal	49993	-23.18	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	25791	24207	1.11	<.0001

Interpretation

- In the 4th table of T-test results, the P value for F test is <.0001. This means that the P value is <0.05, So Null hypothesis will be rejected, implying that the variances are unequal.

- In the 3rd table of T-test results, we check null hypothesis of unequal means with assumption of unequal variances. Here, we find that the t-statistic for unequal variances is -23.18 with a degree of freedom 49993, along with P-value<.0001, inferring that it is not significant.

- As P-value<0.05, Null hypothesis (H_0) is rejected and there is much evidence in support of Alternative Hypothesis(H_a).

Conclusion

- As Null hypothesis (H_0) is rejected from our interpretation, the conclusion is in support of Alternative Hypothesis(H_a).

- Hence, it can be deduced that average weight of baby boys is not equal to average weight of baby girls.

- We also find that average weight of baby boys is greater than average weight of baby girls.

HYPOTHESIS 4

- Compare average weight of babies of smoking mother and non-smoking mother.

Null Hypothesis(H_0):	**Average weight of babies of smoking mother is equal to average weight of babies of non-smoking mother.**
Alternative Hypothesis(H_a):	**Average weight of babies of smoking mother is not equal to average weight of babies of non-smoking mother.**

T-TEST RESULTS

```
title 'T-test on weight by class : momsmoke (if mother smokes or not)';

proc ttest data=work.birth;

class MomSmoke;

var Weight;

run;
```

T-test on weight by class : momsmoke (if mother smokes or not)

The TTEST Procedure

Variable: Weight (Weight)

MomSmoke	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		43467	3402.3	558.0	2.6766	240.0	6350.0
1		6533	3160.9	576.8	7.1358	312.0	5245.0
Diff (1-2)	Pooled		241.5	560.5	7.4376		
Diff (1-2)	Satterthwaite		241.5		7.6213		

MomSmoke	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3402.3	3397.1	3407.6	558.0	554.3	561.8
1		3160.9	3146.9	3174.8	576.8	567.0	586.8
Diff (1-2)	Pooled	241.5	226.9	256.0	560.5	557.1	564.0
Diff (1-2)	Satterthwaite	241.5	226.5	256.4			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	49998	32.46	<.0001
Satterthwaite	Unequal	8474.1	31.68	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	6532	43466	1.07	0.0004

Interpretation

- In the 4th table of T-test results, the P value for F test is <.0001. This means that the P value is <0.05, So Null hypothesis will be rejected, implying that the variances are unequal.

- In the 3rd table of T-test results, we check null hypothesis of unequal means with assumption of unequal variances. Here, we find that the t-statistic for unequal variances is 31.68 with a degree of freedom 8474.1 , along with P-value<.0001, inferring that it is not significant.

- As P-value<0.05, Null hypothesis (H_0) is rejected and there is much evidence in support of Alternative Hypothesis(H_a).

Conclusion

- As Null hypothesis (H_0) is rejected from our interpretation, the conclusion is in support of Alternative Hypothesis(H_a).

- Hence, it can be deduced that average weight of babies of smoking mother is not equal to average weight of babies of non-smoking mother.

- We also find that average weight of babies of mother who smokes is less than that of babies of mother who don’t smoke.

THANK YOU