الأخطاء الشائعة في التحليل الإحصائي وأساليب المعاينة في ميدان التطبيق في كلية الإدارة والاقتصاد

أ.عربية عبد الرحمن داؤد* أ.خوله خالد إسماعيل**

المستخلص

نظراً لكثرة تكرار الأخطاء الشائعة في التحليلات الإحصائية ولسنوات عديدة فقد تم اخذ عينة عشوائية من رسائل الماجستير واطاريح الدكتوراه من مكتبة كلية الإدارة والاقتصاد/جامعة الموصل لأقسام (إدارة الأعمال والإدارة الصناعية والاقتصاد) ولسنوات عديدة وقد تم بيان وعرض هذه الأخطاء وكيفية تجاوزها .

Abstract

For many year, too many common mistakes has been noted in statistics analysis, for this reason a random sample has been taken of Master and PhD thesis from Administration and Economic College Library/Mosul university, for departments of Business Management, Industrial Management and Economic.

These mistakes are demonstrated and then we show how we can deal with these mistakes.

المقدمة :-

تعد رسائل الماجستير والدكتوراه كمصادر أساسية لطلبة الدراسات الأولية والعليا . ونظراً لأهمية التحليل الإحصائي وضرورة صحة استخدامه وذلك للحصول على نتائج منطقية فقد تم اخذ عينة عشوائية من هذه الرسائل والاطاريح من كلية الادارة والاقتصاد /جامعة الموصل ، وتم التركيز على الجانب العملي والإحصائي لها وبيان الأخطاء التي يمكن أن يقع بها طالب الدراسات العليا من وضع الفرضية ولحين اتخاذ القرار . ونظراً لتكرار هذه الأخطاء وبشكل تتابعي منذ سنوات عديدة مما جعل هذه الأخطاء وكأنها قاعدة أساسية صحيحة يُستند عليها في الرسائل والبحوث المستقبلية .

وتعتبر هذه مشكلة كبيرة جداً لا يمكن السكوت عليها وذلك لأنها تنافي النظريات الإحصائية وتبدو كأنها نظريات صحيحة يمكن الاعتماد عليها بشكل كبير (وخاصة من قبل الباحثين في اختصاصات الإدارة والاقتصاد)

وقد تم عرض بعض هذه الأخطاء في بحثنا هذا مع تبيان الطرق الصحيحة الواجب استخدامها في التحليلات الإحصائية لهذه الرسائل والاطاريح علماً بأن هذه الطرق ليست بدائل ولكنها تصحيح لما هو وارد في هذه الرسائل.

مشكلة البحث :

وجود أخطاء في الجانب العملي (الإحصائي) في رسائل الماجستير واطاريح الدكتوراه تتراوح بنسبة أكثر من 95% من الرسائل الموجودة في مكتبة كلية الإدارة والاقتصاد/جامعة الموصل .

* جامعة الموصل /كلية الادارة والاقتصاد

** جامعة الموصل /كلية الادارة والاقتصاد

هدف البحث :

نظراً لكبر حجم الأخطاء الإحصائية والاعتماد عليها بشكل تتابعي وخلال سنوات عديدة مما جعل هذه الأخطاء كقاعدة أساسية يعتمد عليها الطالب أو الباحث بدون الرجوع إلى المصادر العلمية الصحيحة والموثوقة من الناحية الإحصائية فقد تم عرض هذه الأخطاء وبيان كيفية تصحيحها وتجاوزها وذلك بذكر بعض المصادر التي يمكن إن يعتمد عليها وتصحيح ما هو سائد في الجانب العملي للرسائل المستقبلية .

عرض الأخطاء :

بالرغم من دراسة مادة الإحصاء المتقدم كمادة أساسية في مرحلة الماجستير والدكتوراه بالإضافة إلى مبادئ الإحصاء لطلبة الدراسات الأولية ولجميع أقسام الكلية إلاّ أن معظم الرسائل والاطاريح لا تزال في موضع شك وينبثق الشك مما يلي :-

1- تحديد حجم العينة.[4]

إن طبيعة المشكلة المعروضة للبحث والإحاطة بتصنيفها تقتضي التحري عن نوع التصنيف لتحديد حجم العينة. إذ إن حجم العينة ليس هدفاً مجرداً عن نوع المشكلة ، فالحجم المعتمد يهدف الى تقدير نسبة غير معروفه للمجتمع لظاهرة ما (p) هو غير الحجم الذي نعتمده لتقدير المتوسط (μ) الغير معروف للمجتمع لظاهرة أخرى .

ومن العينات الأكثر تطبيقاً في ميدان العمل :

أ- العينة العشوائية البسيطة

ب- العينة العشوائية الطبقية

أ) **العينة العشوائية البسيطة :** يتحدد حجم العينة العشوائية البسيطة اعتماداً على مؤشر النسبة .

فإذا اختيرت عينة من مجتمع وقدرت لها النسبة التي يرمز لها بالرمز p وكان توزيع المعاينة هو التوزيع الطبيعي التقريبي بوسط

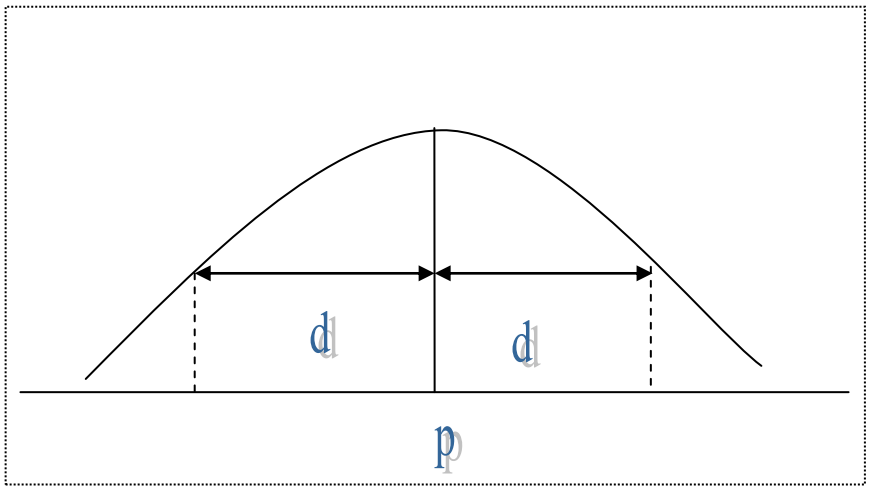
$$E(\hat{P}) = P$$

وتباين

$$\text{var}(\hat{P}) = \frac{p(1-p)}{n}$$

أي بمعنى $V(\hat{P}) = \sigma^2$ ، فأن التعبير عن توزيع المعاينة هو كما في الشكل (1).

شكل(1)
توزيع المعاينة للنسب



$\hat{p} + Z_{1-\alpha/2}\sqrt{p(1-p)/n}$ $\hat{p} - Z_{1-\alpha/2}\sqrt{p(1-p)/n}$

تمثل القيمة (d) المسافة (الفرق) عن نسبة المجتمع ، ويمكن التعبير عنها بالصيغة التالية

$$d = Z_{1-\alpha/2}\sqrt{p(1-p)/n} \quad (1) \qquad \text{.....}$$

بعبارة أخرى إن الحد الأقصى للفرق المطلق بين القيمة الحقيقة للنسبة والقيمة التقديرية (المسافة) كما في الشكل (1) اعلاه. d

ويقدر حجم العينة من العلاقة (1) بحل تلك المعادلة بدلالة n وكما يأتي

$$n = \frac{Z^2_{1-\alpha/2}P(1-P)}{d^2} \qquad \text{.....} \quad (2)$$

علماً أن اكبر حجم للعينة التي نختارها لتقدير النسبة هو عندما تكون النسبة في المجتمع مساوية إلى (0.5) واعتماداً على الصيغة رقم (2) يتم حساب حجم العينة في ضوء المعايير الأساسية الثلاثة التالية

1-درجة الثقة المرغوب فيها : هي القيم الجدولية لـ Z

2-الخطأ المسموح به (d)

3-النسبة المطلوب تقديرها (p)

وقد تم اعتماد جداول قياسية في ضوء المعايير الثلاثة المذكورة لقيم مختارة لكل من درجة الثقة المعبر عنها بقيمة Z الجدولية المقابلة لتلك الدرجة ، الخطأ المطلق المسموح به (d) ، النسبة المتوقعة الحقيقية للمجتمع (p) كما مبين في الجدول رقم (1) أدناه:

جدول رقم (1) حجم العينة اللازم لتقدير نسبة p ضمن الفرق المطلق المسموم به (d)

أ. حدود الثقة (C=99%)

	نسبة المجتمع P																		
d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	3152	5972	8461	10617	12442	13935	15096	15926	16424	16589	16424	15926	15096	13935	12442	10617	8461	5972	3152
0.02	788	1493	2115	2654	3111	3484	3774	3981	4106	4147	4106	3981	3774	3484	3111	2654	2115	1493	788
0.03	350	664	940	1180	1382	1548	1677	1770	1825	1843	1825	1770	1677	1548	1382	1180	940	664	350
0.04	197	373	529	664	778	871	944	995	1026	1037	1026	995	944	871	778	664	529	373	197
0.05	126	239	338	425	498	557	604	637	657	664	657	637	604	557	498	425	338	239	126
0.06	88	166	235	295	346	387	419	442	456	461	456	442	419	387	346	295	235	166	88
0.07	64	122	173	217	254	284	308	325	335	339	335	325	308	284	254	217	173	122	64
0.08	49	93	132	166	194	218	236	249	257	259	257	249	236	218	194	166	132	93	49
0.09	39	74	104	131	154	172	186	197	203	205	203	197	186	172	154	131	104	74	39
0.10	32	60	85	106	124	139	151	159	164	166	164	159	151	139	124	106	85	60	32
0.11	26	49	70	88	103	115	125	132	136	137	136	132	125	115	103	88	70	49	26
0,12	22	41	59	74	86	97	105	111	114	115	114	111	105	97	86	74	59	41	22
0.13	19	35	50	63	74	82	89	94	97	98	97	94	89	82	74	63	50	35	19
0.14	16	30	43	54	63	71	77	81	84	85	84	81	77	71	63	54	43	30	16
0.15	14	27	38	47	55	62	67	71	73	74	73	71	67	62	55	47	38	27	14
0.20	8	15	21	27	31	35	38	40	41	41	41	40	38	35	31	27	21	15	8
0.25	5	10	14	17	20	22	24	25	26	27	26	25	24	22	20	17	14	10	5

تابع جدول رقم (1)
ب- حدود الثقة (C=95%)

نسبة المجتمع P																			
d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	1825	3457	4898	6147	7203	8067	8740	9220	9508	9604	9508	9220	8740	8067	7203	6147	4898	3457	1825
0.02	456	864	1225	1537	1801	2017	2185	2305	2377	2401	2377	2305	2185	2017	1801	1537	1225	864	456
0.03	203	384	544	683	800	896	971	1024	1056	1067	1056	1024	971	896	800	683	544	384	203
0.04	114	216	306	384	450	504	546	576	594	600	594	576	546	504	450	384	306	216	114
0.05	73	138	196	246	288	323	350	369	380	384	380	369	350	323	288	246	196	138	73
0.06	51	96	136	171	200	224	243	256	264	267	264	256	243	224	200	171	136	96	51
0.07	37	71	100	125	147	165	178	188	194	196	194	188	178	165	147	125	100	71	37
0.08	29	54	77	96	113	126	137	144	149	150	149	144	137	126	113	96	77	54	29
0.09	23	43	60	76	89	100	108	114	117	119	117	114	108	100	89	76	60	43	23
0.10	18	35	49	61	72	81	87	92	95	96	95	92	87	81	72	61	49	35	18
0.11	15	29	40	51	60	67	72	76	79	79	79	76	72	67	60	51	40	29	15
0.12	13	24	34	43	50	56	61	64	66	67	66	64	61	56	50	43	34	24	13
0.13	11	20	29	36	43	48	52	55	56	57	56	55	52	48	43	36	29	20	11
0.14	9	18	25	31	37	40	45	47	49	49	49	47	45	41	37	31	25	18	9
0.15	8	15	22	27	32	36	39	41	42	43	42	41	39	36	32	27	22	15	8
0.20	5	9	12	15	18	20	22	23	24	24	24	23	22	20	18	15	12	9	5
0.25	*	6	8	10	12	13	14	15	15	15	15	15	14	13	12	10	8	6	*

تابع جدول رقم (1)
ج- حدود الثقة (C=90%)

نسبة المجتمع P																			
d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	1285	2435	3450	4330	5074	5683	6156	6494	6697	6765	6697	6494	6156	5683	5074	4330	3450	2435	1285
0.02	321	609	863	1082	1268	1421	1539	1624	1674	1691	1974	1624	1539	1421	1268	1082	863	609	321
0.03	143	271	383	481	564	631	684	722	744	752	744	722	684	631	564	481	383	271	143
0.04	80	152	216	271	317	355	385	406	419	423	419	406	385	355	317	271	216	152	80
0.05	51	97	138	173	203	227	246	260	268	271	268	260	246	227	203	173	138	97	51
0.06	36	68	96	120	141	158	171	180	186	188	186	180	171	158	141	120	96	68	36
0.07	26	50	70	88	104	116	126	133	137	138	137	133	126	116	104	88	70	50	26
0.08	20	38	54	68	79	89	96	101	105	106	105	101	96	89	79	68	54	38	20
0.09	16	30	43	53	63	70	76	80	83	84	83	80	76	70	63	53	43	30	16
0.10	13	24	35	43	51	57	62	65	67	68	67	65	62	57	51	43	35	24	13
0.11	11	20	29	36	42	47	51	54	55	56	55	54	51	47	42	36	29	20	11
0.12	9	17	24	30	35	39	43	45	47	47	47	45	43	39	35	30	24	17	9
0.13	8	14	20	26	30	34	36	38	40	40	40	38	36	34	30	26	20	14	8
0.14	7	12	18	22	26	29	31	33	34	35	34	33	31	29	26	22	18	12	7
0.15	6	11	15	19	23	25	27	29	30	30	30	29	27	25	23	19	15	11	6
0.20	*	6	9	11	13	14	15	16	17	17	17	16	15	14	13	11	9	6	*
0.25	*	*	6	7	8	9	10	10	11	11	11	10	10	9	8	7	6	*	*

تشير العلامة (*) الى ان حجم العينة اللازم اقل من 5 وحدات

ب- تحديد حجم العينة في المعاينة العشوائية الطبقية

لوحظ انه عند تحديد حجم العينة في المعاينة العشوائية البسيطة إن (n) تعتمد على الخطأ المعياري للمتوسط ، إما إذا تطلبت حاجة المسح استخدام أسلوب المعاينة العشوائية الطبقية ، فأن الحاجة تقتضي توفير قيمة الخطأ المعياري لتقدير العينة الطبقية .لكن أسلوب المعاينة الطبقية يرتبط بأربع طرق رئيسة لتوزيع حجم العينة هي التوزيع المتساوي ، التوزيع المتناسب، التوزيع الأمثل ، توزيع نيمان. وان لكل طريقة صيغة معينة للخطأ المعياري للتقدير .

إن صيغة تحديد حجم العينة تنطلق من العلاقة بين الخطأ المسموح به (d) والخطأ المعياري للتقدير حيث أن :

$$d = ZS(\bar{x})$$

فإذا حددنا قيمة الخطأ (d) ومستوى الثقة (Z)فأن بألامكان اعتماد الصيغة الآتية للخطأ المعياري للتقدير للعينة الطبقية :

$$S^2(\bar{x}_{st}) = \frac{d^2}{z^2} = B^2 \text{.....(3)}$$

ومعروف أن تباين التقدير يقل مع زيادة حجم العينة ، ولذالك يمكن الوصول إلى حجم العينة (n) الذي يعطي قيمة للتباين مساوية إلى(B^2) التي حددت مسبقا بموجب العلاقة (3) أي أن تحديد حجم العينة ينطلق من قاعدة مساواةB^2 وتباين التقدير .لكن تباين التقدير ، كما ذكرنا ، يختلف باختلاف طرق توزيع حجم العينة في المعاينة الطبقية ، ويأخذ الصيغ الآتية لطرق التوزيع المتساوي والمتناسب والأمثل وتوزيع نيمان على التوالي:

$$S^2(\bar{x}_{eq}) = \frac{L}{N^2}\sum_{h=1}^{L}\frac{N_h{}^2 S_h{}^2}{n} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h{}^2$$

$$S^2(\bar{x}_{prop}) = \frac{1}{N}\sum_{h=1}^{L}\frac{N_h S_h{}^2}{n} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h{}^2$$

$$S^2(\bar{x}_{opt}) = \frac{1}{N^2}\frac{1}{n}\left[\sum_{h=1}^{L} N_h S_h \sqrt{C_h}\right]\left[\sum_{h=1}^{L}\frac{N_h S_h}{\sqrt{C_h}}\right] - \frac{1}{N^2}\sum N_h S_h{}^2$$

$$S^2(\bar{x}_{Ney}) = \frac{1}{N^2}\frac{\left(\sum_{h=1}^{L} N_h S_h\right)^2}{n} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h{}^2$$

تحديد حجم العينة

في حالة التوزيع المتساوي ، باستخدام العلاقة B والخطأ المعياري ، حيث أن

$$B^2 = S^2(\bar{x}_{eq}) = \frac{1}{N^2}\sum_{h=1}^{L}\frac{N_h{}^2 S_h{}^2}{n} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h{}^2 \quad \text{.....زز.....(4)}$$

وبعد التبسيط تكون

$$N^2 B^2 + \sum_{h=1}^{L} N_h S_h{}^2 = \frac{1}{n} L \sum_{h=1}^{L} N_h{}^2 S_h{}^2 \quad \text{.....(5)}$$

وبحل المعادلة بدلالة (n) يكون حجم العينة في حالة التوزيع المتساوي كالأتي:

$$n_{eq} = \frac{L\sum_{h=1}^{L} N_h{}^2 S^2{}_h}{N^2 B^2 + \sum_{h=1}^{L} N_h S^2{}_h}$$

إما في حالات التوزيع المتناسب والأمثل فتستبدل قيمة $S^2(\bar{x}_{eq})$ **في العلاقة (4)بقيم** $S^2(\bar{x}_{prop})$ **و** $S^2(\bar{x}_{opt})$ **و** $S^2(\bar{x}_{Neym})$ **على التوالي ، وبعد التبسيط يكون حجم العينة للطرق الثلاثة كالآتي :**

$$n_{prop} = \frac{N\sum_{h=1}^{L} N_h S_h{}^2}{N^2 B^2 + \sum_{h=1}^{L} N_h S_h{}^2} \quad \text{.....(6)}$$

$$n_{opt} = \frac{\left(\sum N_h S_h \sqrt{C_h}\right)\left(\sum N_h S_h \Big/ \sqrt{C_h}\right)}{N^2 B^2 + \sum_{h=1}^{L} N_h S_h{}^2} \quad \text{.....(7)}$$

$$n_{Ney} = \frac{\left(\sum N_h S_h\right)^2}{N^2 B^2 + \sum N_h S_h{}^2} \quad \text{.....(8)}$$

وبعد تقدير حجم العينة الكلي يوزع على الطبقات بموجب إحدى الطرق الأربعة المذكورة .

ملاحظات مهمة:

أ- إذا كانت قيمة معامل تصحيح المجتمع المحدود قريبة من الواحد (fpc=1) فأن مقام المعادلات(5,6,7,8) يضم الحد $N^2 B^2$ **فقط .**

ب- أن جميع الصيغ المذكورة(5,6,7,8) مبنية على أساس أن قيم تباينات كل طبقة في المجتمع S_h معلومة . لكن هذا أمر نادر في الحياة العملية ، لذلك تقدر قيمة S_h إما من عينات مناسبة ، أو من مصادر أخرى .

2- استمارة الاستبيان

جميع الباحثين يقومون باعداد استمارة استبيان بوضع أسئلة خاصة وشخصية لأفراد عينة الدراسة ولم يتم تحليل هذه البيانات الخاصة بهذه الأسئلة وبيان تأثيرها المباشر على الدراسة علماً بأن لها تأثير مباشر. ويمكن أن تدخل ضمن تحليل العوامل المؤثرة على المشكلة قيد الدراسة مثل (الجنس ، العمر ، المنصب ، التحصيل الدراسي ، ..... الخ).

ويمكن تصنيف البيانات الى نوعين منها الكمية وهي البيانات القابلة للقياس كالوزن والطول وغيره ،والبيانات النوعية وهي البيانات الغير قابلة للقياس كاللون والصنف وغيرها ،الا انه بالامكان وصفها وتحويلها الى قياس رقمي ومنها الاصناف الاسمية (nominal gatogaries) والاصناف المرئية (order gatogaries) فالاصناف الاسمية هي التي لا تشترط الترتيب ، اما الصناف المرئية فان اساس تحليلها يمثل بمدى استجابة المشاهدة او العينة لصيغة من مقياس مقدر ،حيث يعرف المقياس بانه عبارة عن البعد المرتب الذي يتطلب منه اتخاذ قرار حول الكمية او الكثافة او الدرجة ،وتوجد عدة انواع من المقاييس المقدرة منها مقياس ليكرت.

3- مقياس ليكرت.[6]

استخدم هذا المقياس من قبل معظم طلاب الدراسات العليا في إعداد الرسائل والاطاريح الخاصة بهم في كلية الإدارة والاقتصاد عامة وقسمي (إدارة الإعمال ، الإدارة الصناعية) خاصة .

وهذا المقياس هو أسلوب لقياس السلوكيات يستعمل في الاختبارات النفسية والذي استنبطه عالم النفس الانكليزي (رينميس ليكرت) يستعمل في الاستبيانات وبخاصة في مجال الإحصاءات . ويعتمد المقياس على ردود تدل على درجة الموافقة أو الاعتراض على صيغة ما .

ويمكن تعريفه بشكل دقيق بكونه مجموعة الإجابات المحصلة حول (فقرات ليكرت)، إما فقرات ليكرت فتتألف من قسمين:

أ- الجذع : وهي جملة تحدد سلوكية ما .

ب- السلم : وهو مقياس يستعمل لتحديد درجة الموافقة والاختلاف مع جملة الجذع .

وبعد الانتهاء من الإجابة على جميع الأسئلة ، يمكن تحليل كل جذع على حدى كما يمكن جمع إجابات مجموعات من الجذع للحصول على نتيجة جماعية لكل مجموعة ويمكن تحليل الناتج بناءاً على طريقتين الأولى اعتبارها قيم تراتبية والثانية اعتبارها قيم فترات (interval – level) وهناك اختلاف على تحديد إي من الطريقتين أجدى، عند اعتماد نتيجة اختيار ليكرت كقيمة تراتبيه تمثل النتيجة كرسم بياني شريطي (Bar-chart) وملخص متوسطه يكون الوسيط أو المنوال وليس المعدل .

ويقاس مقياس النزعة المركزية من خلال الامتداد عبر الأرباع (وليس من خلال الانحراف المعياري) كما يمكن تحليلها باستخدام الاختبارات اللامعلمية فالإحصاء اللامعلمي (Non Parametric Statistics) هو الإحصاء الذي لا يتقيد بالشروط التي يتقيد بها الباحث عند استخدامه الإحصاء المعلمي خاصة فيما يتعلق بحجم العينة وبشكل التوزيع التكراري ،وغالبا ما يستخدم في العينات الصغيرة وفي التوزيعات الحرة غير المقيدة بالتوزيع الطبيعي .ويفترض الإحصاء اللامعلمي أن تكون مشاهدات العينة مستقلة وان المتغيرات مستمرة .

ومن الاختبارات المستخدمة في التوزيع اللامعلمي :[2]

أ-اختبار ويلكوكس (اختبارات اشارات الرتب)Wilcoxon Signed –Ranks Test

ويستخدم هذا الاختبار عندما تكون العينة مسحوبة من مجتمع يتوزع توزيعا مستمرا مجهول الوسيط وان عينات المجتمع متماثلة ومشاهدات العينة مستقلة فيما بينها .

ب-اختبار مان ويتني Mann –Whitney Test

ان هذا الاختبار يعوض عن اختبار t للمقارنة بين متوسطي عينتين مستقلتين ويمكن استخدام هذا الاختبار حتى في العينات الصغيرة جدا .

ج-اختبار كروسكل-والسKruskal –Wallis Test

وهي الطريقة البديلة لجدول تحليل التباين والتي يتم فيها اختبار تساوي عدة اوساط حسابية ، وتعتمد الطريقة اساسا على كون المتغيرات تحت الدراسة لها نفس الشكل وان العينات المختلفة مستقلة عن بعضها .

د- اختبار الاستقلال بين المتغيرات Test for Independence

يمكن استخدام مربع كاي عندما يراد اختبار فرضية حول استقلال متغيرين او (ظاهرتين) سواء اكانتا نوعية او وصفية .وعادة تنظم البيانات في جدول مزدوج يسمى جدول التوافق، حيث تمثل الصفوف ظاهرة معينة والاعمدة تمثل الظاهرة الاخرى ، وعند ايجاد قيمة مربع كاي يمكننا الحصول على ثلاثة انواع من المعاملات التي تقيس درجة العلاقة بين المتغيرات :

1- **فاي(Ø)**
2- **معامل الاقتران(C)**
3- **Vكرامر (V)**

كل هذه الاحصاءات هي بسيطة وسهلة الحساب ،وكلها تعتمد على ايجاد مربع كاي في البداية .القيمة المستخرجة لكل احصاءة ستكون دائما بين الصفر والواحد. الارتباط السالب مع هذه الاحصاءات رياضيا مستحيل واستخداماتها تعتمد على حجم المصفوفة وكذلك فيما اذا كان او لم يكن المتغيرات تحت الدراسة لديها نفس العدد من القيم الممكنة .ان فاي تستخدم عندما يكون كلا المتغيرين تحت الدراسة لديهما بالضبط 2 من القيم الممكنة .واذا كان هذا صحيح فان المصفوفة ستكون (2X2).

اما (C)تستخدم عندما تكون 3 أو اكثر من القيم لكل متغير وكذلك تكون بعدد متساوي من القيم ، أي ان المصفوفة ستكون (3X3,4X4,...).

ويمكن ايجادها بالصيغة بالتالية :

4-الفرضية الإحصائية

أن معظم الرسائل والاطاريح الموجودة في مكتبة الكلية لا تزال موضع شك ، وينبثق الشك من وضع الفرضيات الخاطئة في صياغتها .

من المعروف إن الفرضية هي عبارة عن ادعاء أو تصريح قد يكون صائباً أو خطأ حول معلمة أو أكثر لمجتمع أو مجموعه من المجتمعات ، فالإحصائي أو الباحث يحاول دائما أن يضع الفرضيات بشكل يأمل أن يرفضها لان قبول الفرضية لا يعنى بالضرورة كونها صحيحة [1]،[3]،[7]،[8]،[9]. أي عدم وجود أدلة كافية لرفضها من بيانات العينة . إما إذا رفضت الفرضية بناءاً على المعلومات الموجودة في بيانات العينة فان ذلك يعني أن الفرضية خاطئة . وهذا التعريف الذي يستدل به الباحث ، لكونه التعريف الأساس والدليل ، في وضع فرضياته بناءاً على مشكلة يتم تحديدها ومحاولة دراستها ومعالجتها علماً بأن الفرضية الإحصائية هي أساس البحث العلمي بعد تحديد المشكلة والتي تعتبر عنصر أساسي في منهجية البحث حيث يعتمد عليها في جميع التحليلات الإحصائية منها (معامل الارتباط ، الانحدار ، الاستقلالية ، ..... الخ) حيث أن جميع الباحثين يضعون فرضية العدم بصيغة القبول وعندها سيكون قرار معاكس فيحصل على نتيجة خاطئة إضافة إلى ذلك سيكون الجانب التطبيقي برمته غير صحيح من الناحية الاحصاية . لذا نقترح أن تصاغ الفرضية بشكل صحيح بالاعتماد على المصادر الإحصائية المتوفرة والقيام بتحليل الناتج بشكل صحيح أيضا بناءا على الفرضية التي تم وضعها من قبل الباحث.

5- معامل الارتباط .[8]

جميع الباحثين في عينة البحث الذين يستخدمون معامل الارتباط البسيط (بيرسون) في تحليلاتهم في حين أن بياناتهم يجب أن يتم تحليلها بالاعتماد على معامل الارتباط (سبيرمان) لان متغيرات الاستجابة تعتبر متغيرات لا معلميه إضافة إلى أن هناك فرق حسابي واضح في الحصول على النتيجة النهائية بين هذين المقياسين .

كما أن جميع الباحثين لم يتطرقوا إلى اختبار معامل الارتباط اختبارا إحصائيا للتأكد من معنوية معامل الارتباط . هذا وان معظم البرامج الجاهزة المستخدمة في التحليلات الإحصائية تحسب معامل ارتباط (بيرسون) ونختبره كما أن اغلب الباحثين عند توضيح درجة الارتباط بين متغيري الدراسة يشير إلى أن المتغير الأول بالمتغير المعتمد والأخر بالمتغير المستقل وهذا غير صحيح من الناحية المنطقية والإحصائية .

6- معامل الانحدار. [5]

أن تحليل الانحدار يعد من أوسع العلوم الإحصائية استخداماً في مختلف العلوم ذلك انه يحدد بوضوح العلاقة بين المتغيرات على هيئة معادلة ويستدل من تقدير معلماتها على أهمية وقوة واتجاه هذه العلاقة . كما يبين تقدير الاستجابة والتنبؤ بما يفيد كثيرا في التخطيط واتخاذ القرارات الرصينة حولها . ونلاحظ أن معظم الرسائل العلمية والاطارح في كلي الإدارة والاقتصاد تستخدم أسلوب الانحدار في التحليل ولكن يتجاهلون بعض الأمور المهمة منها :

أ) أكثر الباحثين يؤكدون أن هناك عدة أسباب تؤثر على متغير معتمد معين ومن ثم يتم اختيار نموذج يحتوي على متغير مستقل واحد فقط بالاعتماد على معادلة الانحدار البسيط (لسهولة العمل به)

ب) عدم رسم البيانات المتوفرة لدى الباحثين لتحديد نموذج الانحدار الملائم للمشكلة.
ت) معظم الباحثين لم يتطرقوا إلى التداخل الخطي بين البيانات وكيفية معالجته قبل التحليل . لذا سوف يظهر النموذج ذو معامل تحديد R^2 كبير جدا يقترب من 100% وهذا غير منطقي.
ث) عند اختبار النموذج باستخدام و R^2 ,F,وحسب ما يظهر في نتائج الحاسبة الالكترونية فإن الباحث يعلق على R^2 فقط باعتباره المقياس الوحيد لقياس العلاقة بين متغيرات الدراسة وإهمال بقية النتائج مثلاً قيمة B1 , F , B2 .... الخ
ج) معظم الباحثين يضعون نتائج التحليل في جداول من إعداد الباحث والمفروض على الباحث أن يضع النتائج كما تظهر نصاً ومباشرة من الحاسبة الالكترونية ، لكي يتم التعليق والاستفادة من جميع النتائج التي تظهر في هذا التحليل .

7- على جميع الباحثين أن يقوموا برفق بيانات البحث التي اعتمدها ، للتأكد من نتائج التحليل ابتعاداً عن التلاعب بالنتائج حسب ما يتفق مع الجانب النظري في البحث .

التوصيات

بعد عرض الأخطاء الشائعة والمشار إليها سابقاً نوصي بما يلي :

1) تحديد حجم العينة بشكل علمي وصحيح وكما ورد في متن البحث ، كون تحديد حجم العينة بشكل صحيحاً يترتب عليه كل ما يتعلق بالاختبارات الإحصائية .
2) ضرورة إعداد استمارة الاستبيان بشكل صحيح بحيث يمكن للباحث الاستفادة من جميع الأسئلة المطروحة في الاستمارة واستبعاد الأسئلة التي لم يستفد منها الباحث في التحليل الإحصائي.
3) ضرورة استخدام مقياس ليكرت بشكل علمي صحيح كما ورد ذكره متن البحث .
4) وضع الفرضيات الإحصائية بشكل صحيح وهو اهم خطوة في التحليل الإحصائي لأنه سينبثق عنه القرار النهائي .
5) يجب استخدام معامل الارتباط الملائم للبيانات كما ذكرنا في بحثنا مع القيام باختبار معنويته .
6) إذا تم استخدام معادلة الانحدار من قبل الباحث في تحليله يأخذ بنظر الاعتبار:

أ- كل العوامل المؤثرة على المتغير المعتمد كمتغيرات مستقلة .
ب- ضرورة رسم البيانات لتحديد نوع النموذج.
ت- معالجة البيانات التي تتميز بالتداخل الخطي
ث- يجب على الباحث أن يبين تأثير كل من (B1,B2,..... R^2 F,) دون الاقتصار على R^2 فقط
ج- يجب عرض نتائج التحليل كما تظهر في الحاسبة الالكترونية مباشرة دون اللجوء إلى جداول من إعداد الباحث

7) يجب رفق بيانات البحث التي اعتمدها الباحث أولا للتأكد من التحليل وثانياً لكي يتم الاستفادة منها من قبل الباحثين في المستقبل

المصادر العربية:

1. الراوي، خاشع ،1979،"المدخل إلى الإحصاء" / مؤسسة دار الكتب للطباعة والنشر/ جامعة الموصل .
2. القرشي ،إحسان كاظم ، 2007، "الطرائق المعلمية والطرائق اللا معلمية في الاختبارات الإحصائية"،الطبعة الاولى ،
3. المشهداني،محمود ، هرمز ، أمير حنا ،1989،" الإحصاء "، جامعة بغداد ، بيت الحكمة .
4. العلاق ،حمد ، عدنان شهاب ، مهدي محسن ،2001،"أساليب المعاينة في ميدان التطبيق" ، المعهد العربي للتدريب ، البحوث الإحصائية ، بغداد.
5. محمد ، فراس حامد ذنزن ، 2010،"دراسة مقارنة بين متغيرات القيم المطلقة وتقديرات المربعات الصغرى باستخدام البرمجة الخطية "،رسالة ماجستير غير منشورة ،جامعة الموصل،العراق.
6. رينسيس ليكرت ،1932،" أسلوب البحث مقياس السلوكيات (باللغة الانكليزية)" ، أرشيف علم النفس ، ص 145 ،الموقع في الشبكة العالمية.

http://ar.wikipedia.org/wiki

المصادر الانكليزية:

7- Cox,D.R.,2007,"Applied Statistics :a review " ,Nuffield College ,Oxford OX1 1F,UK.

8- Hamburg,M.,1985,"Basic Statistic ;A modern Approach ",Harcourt School;3 Sub Addition.

9-Maureen,Hillenmeyer,2005,"Applied Staistics" ,WWW.Ten.Uoc.gr/~Vagelis/Courses/EM272/Stats_M.Hillen-meyer.pdf.

.....

.....

.....