دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان:	مقارنة لبعض طرائق تقدير ومعالجة القيم المفقودة في تصاميم التجارب الأساسية
المؤلف الرئيسي:	هدبة، وكاع علي
مؤلفين آخرين:	دبدوب، مروان عبدالعزيز(مشرف)
التاريخ الميلادي:	2005
موقع:	الموصل
الصفحات:	1 - 120
رقم MD:	552490
نوع المحتوى:	رسائل جامعية
اللغة:	Arabic
الدرجة العلمية:	رسالة ماجستير
الجامعة:	جامعة الموصل
الكلية:	كلية علوم الحاسبات والرياضيات
الدولة:	العراق
قواعد المعلومات:	Dissertations
مواضيع:	الإحصاء، الأساليب الإحصائية، المعادلات الرياضية
رابط:	http://search.mandumah.com/Record/552490

© 2020 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

وزارة التعليم العالي والبحث العلمي
جامعة الوصل
كلية علوم الحاسبات والرياضيات

مقارنة لبعض طرائق تقدير ومعالجة القيم المفقودة في تصاميم التجارب الاساسية

رسالة مقدمة

إلى

مجلس كلية علوم الحاسبات والرياضيات في جامعة الموصل
وهي جزء من متطلبات نيل شهادة الماجستير في علوم الإحصاء

من قبل الطالب

وكاع علي هدبة

بأشراف

الأستاذ المساعد

مروان عبد العزيز دبدوب

١٤٢٦هـ ٢٠٠٥م

إقرار المشرف

أشهد أن إعداد هذه الرسالة جرى تحت إشرافي في جامعة الموصل وهي جزء من متطلبات نيل شهادة الماجستير في الاحصاء.

التوقيع:
الاسم : أ.م. مروان عبدالعزيز دبدوب
التاريخ: / /

إقرار المقوم اللغوي

أشهد بأن هذه الرسالة الموسومة(**مقارنة لبعض طرائق تقدير ومعالجة القيم المفقودة في تصاميم التجارب الاساسية**) تمت مراجعتها من الناحية اللغوية وتصحيح ما ورد فيها من أخطاء لغوية وتعبيرية وبذلك أصبحت الرسالة مؤهلة للمناقشة بقدر تعلق الأمر بسلامة الأسلوب وصحة التعبير.

التوقيع:
الاسم : د.ميسر حميد مجيد
التاريخ: / /

إقرار رئيس لجنة الدراسات العليا

بناءا" على التوصيات التي تقدم بها المشرف والمقوم اللغوي،أرشح هذه الرسالة للمناقشة.

التوقيع:
الاسم : أ.م.د.طالب شريف جليل
التاريخ: / /

إقرار رئيس قسم الإحصاء

بناءا" على التوصيات التي تقدم بها المشرف والمقوم اللغوي ورئيس لجنة الدراسات العليا،أرشح هذه الرسالة للمناقشة.

التوقيع:
الاسم : أ.م.د.حسن محمد الياس
التاريخ: / /

بسم الله الرحمن الرحيم

﴿اللَّهُ نُورُ السَّمَاوَاتِ وَالأَرْضِ مَثَلُ نُورِهِ كَمِشْكَاةٍ فِيهَا
مِصْبَاحٌ الْمِصْبَاحُ فِي زُجَاجَةٍ الزُّجَاجَةُ كَأَنَّهَا كَوْكَبٌ دُرِّيٌّ
يُوقَدُ مِن شَجَرَةٍ مُّبَارَكَةٍ زَيْتُونِةٍ لا شَرْقِيَّةٍ وَلا غَرْبِيَّةٍ يَكَادُ زَيْتُهَا
يُضِيءُ وَلَوْ لَمْ تَمْسَسْهُ نَارٌ نُّورٌ عَلَى نُورٍ يَهْدِي اللَّهُ لِنُورِهِ مَن
يَشَاء وَيَضْرِبُ اللَّهُ الأَمْثَالَ لِلنَّاسِ وَاللَّهُ بِكُلِّ شَيْءٍ عَلِيمٌ ﴾

صدق الله العظيم

سورة النور

الآية {٣٥}

الإهداء

إلى ...

رسول البشرية نبينا محمد (صلى الله عليه وسلم) .

أخي المفقود دعاءً بالعودة والسلامة .

روح أبي عطر الله ثراه ترحماً .

والدتي الحنونة برا وخفض جناح.

زوجتي شريكة حياتي حباً و وفاءً .

أولادي واخوتي سندي في الحياة عطفا وحنانا .

جميع الذين يتمسّكون بالقيم والذين يبحثون عنها في زمن فقدت فيه القيم .

اهدي جهدي المتواضع .

الباحث

شكر وتقدير

الحمد لله والصلاة والسلام على من لانبي من بعده ، وبعد ...

اقدم شكري وامتناني الى استاذي الفاضل مروان عبد العزيز دبدوب لقبوله تحمل عناء الإشراف على رسالتي ولمتابعته لي طول مدة اعدادها ولما قدمه من ملاحظات قيمة وما ابداه من روح علمية مخلصة افادت الرسالة واغنتها ...

واتقدم بالشكر الى الدكتور علي عزيز علي عميد كلية علوم الحاسبات والرياضيات لدعمه المتواصل لطلبة الدراسات العليا .

كما يدعوني واجب العرفان والوفاء أن اتقدم بشكري وامتناني الى الدكتور حسن محمد الياس رئيس قسم الإحصاء والى جميع أساتذه القسم كافة لما قدموه لي من عون ومساعده كما اتقدم بالشكر الجزيل الى الدكتور خالد محمد داود لمساعدته لي في توفير البيانات الإحصائية .

واقدم شكري الى جميع زملائي طلبة الدراسات العليا والأخ مسؤول مختبر الإحصاء صفوان ناظم والى موظفي المكتبة والى كل من مد لي يد المساعدة طول مدة الدراسة أسأل الله أن يوفق الجميع لما يحبه ويرضاه .

الباحث

المحتويات

الموضوع		رقم الصفحة
الفصل الأول : المقدمة والاستعراض المرجعي وهدف البحث		١-٦
١-١	المقدمة	١
١-٢	الاستعراض المرجعي	٣
١-٣	هدف البحث	٦
الفصل الثاني : الجانب النظري		٧-٢٦
٢-١	التعاريف الأساسية في التجارب المصممة	٧
٢-٢	تحليل التباين	٨
٢-٢-١	الفروض الأساسية في تحليل التباين	٨
٢-٢-٢	أسس تصميم التجارب	٩
٢-٣-١	التصميم العشوائي الكامل	٩
٢-٣-٢	تصميم القطاعات العشوائية	١١
٢-٣-٣	تصميم المربع اللاتيني	١٢
٢-٤	القيم المفقودة	١٣
٢-٥	عرض بعض طرائق تقدير القيم المفقودة	١٣
٢-٥-١	طريقة يتز	١٣
٢-٥-٢	طريقة هاري	١٥
٢-٥-٣	طريقة Rubin	٢٣
٢-٥-٤	طريقة هسمان وكلير	٢٤
الفصل الثالث : المتغيرات الوهمية في تحليل الانحدار		٢٧-٤٤
٣-١	المقدمة	٢٧
٣-٢	طرق الترميز	٢٨
٣-٣	النموذج الخطي العام	٣٠
٣-٤	تقدير معلمات نموذج الانحدار المتعدد واختبار الفرضيات	٣١
٣-٥	العلاقة بين تحليل الانحدار وتحليل التباين	٣٤

٣-٦	تحليل الانحدار والترميز الوهمي	٣٥
٣-٦-١	التصميم العشوائي الكامل	٣٥
٣-٦-٢	تصميم القطاعات العشوائية	٣٨
٣-٦-٣	تصميم المربع اللاتيني	٤١
الفصل الرابع :الجانب التطبيقي		٤٥-١١٥
٤-١	المقدمة	٤٥
٤-٢	جمع البيانات	٤٦
٤-٣	تحليل التجربة	٥٠
٤-٣-١	التصميم العشوائي الكامل	٥٠
٤-٣-١-١	في حالة عدم وجود قيم مفقودة	٥٠
٤-٣-١-١-١	طريقة تحليل التباين	٥٠
٤-٣-١-١-٢	طريقة تحليل التباين للإنحدار باستخدام الترميز التأثيري	٥١
٤-٣-١-٢	في حالة وجود قيم مفقودة	٥٢
٤-٣-١-٢-١	في حالة فقدان قيمة واحدة	٥٢
٤-٣-١-٢-٢	في حالة فقدان قيمتين لمعاملات مختلفة	٥٣
٤-٣-١-٢-٣	قيمتين مفقودتين من نفس المعاملة	٥٥
٤-٣-١-٢-٤	ثلاث قيم مفقودة	٥٦
٤-٣-١-٣	المقارنة بين تحليل التباين للإتحدار باستخدام الترميز التأثيري وتحليل التباين للتصميم العشوائي الكامل	٥٨
٤-٣-٢	تصميم القطاعات العشوائي الكاملة	٦٠
٤-٣-٢-١	في حالة عدم وجود قيم مفقودة	٦٠
٤-٣-٢-١-١	طريقة تحليل التباين	٦٠
٤-٣-٢-١-٢	طريقة تحليل التباين للإنحدار باستخدام الترميز التأثيري	٦٠
٤-٣-٢-٢	في حالة وجود قيم مفقودة	٦٢
٤-٣-٢-٢-١	في حالة فقدان قيمة واحدة	٦٢
٤-٣-٢-٢-٢	في حالة فقدان قيمتين لمعاملات مختلفة لنفس القطاع	٦٥
٤-٣-٢-٢-٣	فقدان قيمتين لنفس المعاملة وقطاع مختلف	٧٠
٤-٣-٢-٢-٤	في حالة فقدان ثلاث قيم	٧٥

٤-٣-١-٤	المقارنة بين تحليل التباين للانحدار باستخدام الترميز التأثيري وتحليل التباين في تصميم القطاعات العشوائية الكامل	٨٠
٤-٣-٣	تصميم المربع اللاتيني	٨٤
٤-٣-٣-١	في حالة عدم وجود قيم مفقودة	٨٤
٤-٣-٣-١-١	طريقة تحليل التباين	٨٤
٤-٣-٣-١-٢	طريقة تحليل التباين للإنحدار باستخدام الترميز التأثيري	٨٥
٤-٣-٣-٢	في حالة وجود قيم مفقودة	٨٦
٤-٣-٣-٢-١	في حالة فقدان قيمة واحدة	٨٧
٤-٣-٣-٢-٢	في حالة فقدان قيمتين لنفس العمود ولصفوف ومعاملات مختلفة	٩٠
٤-٣-٣-٢-٣	في حالة فقدان قيمتين في صف وعمود مختلف لنفس المعاملة	٩٤
٤-٣-٣-٢-٤	في حالة فقدان ثلاث قيم	٩٩
٤-٣-١-٥	المقارنة بين تحليل التباين للإنحدار باستخدام الترميز التأثيري وتحليل التباين في تصميم المربع اللاتيني	١٠٤
*	نتائج تحليل التصميم العشوائي الكامل	١٠٩
*	ملخص نتائج الطرائق التقديرية ونتائج تحليل التباين للإنحدار في تصميم القطاعات العشوائية الكامل	١١٠
*	ملخص نتائج الطرائق التقدير ونتائج تحليل التباين للانحدار في تصميم المربع اللاتيني	١١٢
*	الاستنتاجات	١١٤
*	التوصيات	١١٥
*	المصادر	١١٦-١٢٠

1

الفصل الأول

المقدمة والاستعراض المرجعي وهدف البحث

الفصل الأول

المقدمة والاستعراض المرجعي وهدف البحث

١-١ المقدمة :

ان علم تصميم التجارب هو احد فروع علم الإحصاء الذي يختص بتخطيط واستغلال الامكانيات المتاحة للباحث بحيث يمكنه من وضع انسب التصميمات أو الخطط لجمع البيانات وتحليلها والتي يتم الحصول عليها من التجارب وفق أسس علمية تؤدي إلى توضيح نتائج التجارب وتمكن الباحث من اتخاذ القرارات المناسبة بقناعه علمية.

فالتنقيب المستمر عن المعارف والمفاهيم يعني استمرار الاستقصاء عن المعرفة والفهم في سبيل حل مشاكل محدده في جميع أوجه الحياة والتي تربط الحقائق المتباينة التي من خلالها يمكن التنبؤ بأحداث معينة على ان يكون المنهاج المستخدم في حلها معتمداً على اساس الطرق العلمية فحقائق العلم كلها مترابطة وهناك علاقة تجمعها بحيث نقوم بتفسير بعضها من خلال هذا الترابط أو من خلال طريقة التجريب والبحث والملاحظة وهذه العوامل كلها جعلت لعلم تصميم التجارب الأهمية الكبرى في تفسير النتائج والوصول إلى حلول عملية لجميع المشاكل التي تواجه الباحث .

وفي ضوء حركة التطور والتقدم العلمي المعتمد على أجراء التجارب المصممة على اسس علمية دقيقة كان من المهم معالجة بعض المشاكل التي تواجه الباحث عند فقدان بعض وحداته وخاصة في موضوع تصميم وتحليل التجارب وذلك لان البيانات الناقصة لايمكن الاعتماد عليها في تفسير النتائج وتحليلها لما تسببه من نقص في المعلومات والدقة في النتائج التي لها تأثير على اتزان التصميم والخلل في توفر فروض التحليل .

ان مشكلة القيم المفقودة من المشاكل التي لاقت اهتمام الكثير من الباحثين فقد تسقط احياناً مشاهدة من المشاهدات في تجربة ما أو دراسة معينة وعندئذ فان مشاهدة واحدة أو اكثر قد تكون مفقوده في البيانات .

ان فقدان قيمة واحدة أو اكثر في التصميم العشوائي الكامل (Completely Randomized Design) لا يترتب عليه أية مشكلة طالما أن تحليل التباين يمكن اجراؤه في حالة عدم تسأوي عدد المشاهدات إلا ان فقدان اية قيمة في تحليل ذي اتجاهين مثل تصميم القطاعات العشوائية الكاملة (Randomized Complete Block Design) أو تصميم المربع اللاتيني (Latin Square Design) يعني فقدان التعامدية ولحل هذه المشكلة تم احلال قيم تقديريه بدل القيم المفقودة بطرق عديدة كل منها يسعى لاعطاء افضل تقدير .

ومما لا شك فيه بالامكان اعادة التجربة لغرض الحصول على بيانات لا تتضمن قيماً مفقودة ولكن هذا يتطلب مالاً ووقتاً وجهداً بالإضافة الى ذلك فإنه من الممكن ان لا تتوفر نفس ظروف التجربة التي أجريت في المرة الأولى كما قد تظهر قيم مفقودة بعدد اكثر من القيم للتجربة السابقة وقد تحرّف النتائج عن حقيقتها بشكل اكبر .

لقد تم القاء الضوء في هذه الرسالة على موضوع القيم المفقودة ونوقشت بعض طرق التقدير لها وذلك من خلال طرق خاصة بتحليل تصاميم التجارب التي تسعى الى جعل مجموع مربعات الأخطاء أقل ما يمكن أو بالاستعانة بتحليل الإنحدار وخاصة الإنحدار الجزئي باستخدام المتغيرات الوهمية ومحأولة المقارنة بين هذه الطرائق للوصول الى افضل تقدير للقيم المفقودة وبأسهل الطرق .

ولأهمية التصاميم الأساسية الثلاث التصميم العشوائي الكامل (CRD) وتصميم القطاعات العشوائية الكاملة(RCBD) وتصميم المربع اللاتيني (LSD) في كثير من التجارب واعتماد الكثير من التصاميم الاخرى عليها تم دراسة تأثير القيم المفقودة والمقارنة بين طرق تقديرها على هذه التصاميم الثلاث .

١-٢ الاستعراض المرجعي :

لدراسة اية مشكلة والعمل على حلها لابد من اتباع الطريقة العلمية التي بوساطتها يمكن الوصول الى نتائج دقيقة وباستخدام التصاميم التجريبية يمكن الحصول عل نتائج تؤدي الى تقليل قيمة الخطأ التجريبي .

ان فقدان قيمة واحدة أو اكثر يسبب الخلل في اتزان التصميم ومن ثم في شروط التحليل ونتائجه ودقة المعلومات مما يؤدي الى الحصول على مؤشرات ومردودات سلبية لايمكن الاعتماد عليها في التخطيط المستقبلي .

لقد قام العديد من الباحثين بمعالجة مشكلة تقدير القيم المفقودة Missing Values بأساليب وطرق مختلفة.

وفيما يلي استعراض لعدد من طرق تقدير القيم المفقودة والدراسات والبحوث المتعلقة بهذا الموضوع اذ بدأت معالجة مشكلة تقدير القيم المفقودة منذ عام 1930 عندما قام Allan and Wishart بمعالجة مشكلة تقدير القيم المفقودة في تصميمي القطاعات العشوائية الكاملة (RCBD) وتصميم المربع اللاتيني (LSD) ووضعا صيغ مناسبة لمعالجة قيمة مفقودة واحدة وذلك لجعل مجموع مربعات الخطأ اقل ما يمكن .

في عام1933 قدم Yatesصيغتين لتقدير اكثر من قيمة مفقودة واحدة حيث اقترح اسلوب التعويض المتتالي Iterative procedure.

وتمكن Yates عام١٩٣٩ من معالجة اثنين أو اكثر من الصفوف والاعمدة المفقودة عند تحليل تصميم المربع اللاتيني .

وفي عام 1940 استخدم Nair تحليل التغاير بطريقة Bartlet في حالة تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني .

وفي عام 1946 وضح Delury طريقة المعادلات الجبرية في معالجة قيمة مفقودة واحدة أو اكثر في تصميم المربع اللاتيني .

في عام 1950 وضح Smith كيفية استخدام المتغيرات الوهمية في تصميم المربع اللاتيني وكيفة استخدام التغاير في حالة فقدان اكثر من قيمة واحدة.

في عام 1956 عرض Healy and Westmacoot طريقة استخدام الحاسب الالي باستخدام الوسط الحسابي كقيمة تخمينيه ويتم اجراء التحليل بطرح قيمة الخطأ من القيمة التخمينيه ويستمر بالعملية الى ان يكون الخطأ صغيراً جداً.

في عام 1957 عالج Coons مشكلة تقدير القيم المفقودة في تصميم المربع اللاتيني وتصميم القطاعات العشوائية الكاملة عن طريق تحليل التغاير باستخدام المتغيرات الوهمية بالاعتماد على قيمة تقديرية مستخرجه من معامل الإنحدار $\hat{B}$.

واستخدم أسلوب المصفوفات لتقدير القيم المفقودة في تصميم القطاعات العشوائية الكاملة عام 1958 من قبل Glenn and Kramer كما عولجت مشكلة تقدير القيم المفقودة في تصميم المربع اللاتيني عام 1960 من قبل Kramer and Glass في حالة وجود (n) من القيم المفقودة في اغلب التصاميم وذلك بافتراض قيمة تقديرية أولية بدل القيمة المفقودة وأوجد خطأ الخلايا المفقودة وحصل على قيمة تقديرية للقيمة المفقودة.

وفي عام 1964 قام Harry بإيجاد تقدير للقيم المفقودة في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني من خلال حل عدد من المعادلات بعدد القيم المفقودة.

في عام 1972 استخدم Rubin الأسلوب غير المتتالي (non-Itrative) لتقدير القيم المفقودة في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني باستخدام طريقة المربعات الصغرى (Least square) وذلك ليجعل مجموع مربعات خطأ الخلايا مساويا للصفر.

في عام 1973 قام Haseman and Gaylar بمعالجة مشكلة تقدير القيم باستخدام اسلوب (non-Itreative) في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني .

وفي عام 1975 قاما John and Prescott بمعالجة مشكلة تقدير القيم المفقودة باستخدام المصفوفات والمتغيرات الوهمية.

وفي عام 1977 قام Seber باستخدام معكوس المصفوفة معتمداً على طريقة Yates لايجاد القيم المفقودة في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني

وفي عام 1979 استخدم Little تقديرات الإمكان الأعظم في ايجاد البيانات غير الكاملة (الناقصة) .

واستخدم Hunt & Triggs عام 1989 الطرائق التكرارية لتقدير القيم المفقودة باستخدام المصفوفات معتمداً على فكرة Healy (1956) وعممها في معظم التصاميم .

وفي عام 1990 قام Hunt and Bell باستخدام مختلف طرق التعويض المتتالي لتقديرات المربعات الصغرى للقيم المفقودة في مختلف التصاميم .

وفي عام (1995) استخدم Hunt صيغة خاصة لتخمين كفاءة النتائج في حالة فقدان (m) من المشاهدات المفقودة وفي نفس العام استخدم Kirk طريقة نموذج وسط الخلية لايجاد القيم المفقودة في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني .

وفي عام 1997 استخدم Yandell طريقة لايجاد القيم المفقودة في حالة فقدان عدد من مستويات أحد العوامل أو جزء من مستويات العامل أو احدى الخلايا لتصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني مستخدما أوساط الخلايا .

وفي عام 2002 قدم Howell بحثاً بعنوان معالجة مشكلة القيم المفقودة واستخدم فيه عدة طرق منها استخدام المتغيرات الوهمية ودالة الامكان الاعظم لايجاد القيم المفقودة .

ويجب ان لا ننس ما قام به الباحثون العراقيون من خلال المساهمة في تطـوير العلـم فـي مختلف المجالات ووضع افكار جيدة حول علم تصميم التجارب وخاصة في معالجـة بعـض المشاكل التي تلاقي الباحثين ومنها مشكلة القيم المفقودة بطرق عديدة ووضع الحلول المناسبة من اجل الوصول الى نتائج دقيقة ومقبولة وهناك الكثير من الباحثين العراقيين الذين سـاهمو في هذا المجال .

ففي عام 1979 قام الباحث يونان باستخدام طريقـة Haseman and Gaylar وطريقـة Rubin وطريقتين من طرق تحليل التغاير لايجاد القيم المفقـودة وهمـا Shearer وطريقـة Wilkinsom .

في عام 1981 قامت الباحثة يعقوب باستخدام اسلوب تحليل التغاير في التجارب الحقليـة في حالة احتواء البيانات على مشاهدات مفقودة .

وفي عام 1989 استخدم يوسف طريقة Coons وطريقة Rubin وطريقة Haseman and Gaylar وطريقة Shearer لتقدير القيم المفقودة في تصميم القطاعات العشوائية والمربع اللاتيني .

وفي عام 1998قدمت الجبوري دراسة تحليل للقيم الشاذة والمفقودة في تصـميم المربـع اللاتيني والتصميم العشوائي الكامل في حالة تكرار مشاهدات العينة معتمـداً علـى طرائـق مقترحة من طريقة Coons.

وفي عام ٢٠٠١ قام السنجاري باستخدام طريقـة Yates وطريقـة Coons وطريقـة Neter لايجاد القيم المفقودة في تصميم العبور البسيط باسـتخدام المحاكـاة والمقارنـة بـين الطرائق الثلاث .

وفي عام٢٠٠٢ قامت الدأوودي بدراسة بعض طرق تقدير القيم المفقـودة فـي تصـميم القطاعات العشوائية الكاملة وما دام هناك قيم مفقودة في الدراسات والبحـوث وخاصـة فـي المجال الزراعي لما لهذا المجال من ارتباط بالظروف الجوية والطبيعية التي تسـبب غالبـاً تدمير بعض المشاهدات لذا سيبقى موضوع البحث عن افضل الطرق لتقدير معالجة تلك القيم المفقودة قائما .

١-٣ هدف البحث :

ان هدف هذه الرسالة هو الحصول على نتائج دقيقة من خلال تقليل متوسط مربع الخطأ (Mse) الذي يعد احد المعايير المهمة في قياس دقة أية تجربة ولتقليل الجهد والكلفة والوقت في تقدير القيم المفقودة عند فقدان اية قيمة من المشاهدات وهناك اربع طرائق لتقدير قيم المشاهدات المفقودة وهي (Haseman and Gaylar, Rubin , Harry , Yates) وسنستخدم طريقة تحليل الانحدار باسلوب المتغيرات الوهمية و التي لا نرى حاجة لتقدير القيم المفقودة خاصةً اذا كان حجم البيانات كبيراً وكانت المتغيرات كثيرة وفيها قيم مفقودة حيث توفر برامج الحاسبة في الوقت الحاضر يسهل استخدام هذه الطريقة ومن الاهداف .

أ-استخدام اسلوب تحليل الإنحدار عن طريق الترميز التاثيري للحصول على جدول تحليل التباين لكل من التصاميم الأساسية في حالة :-

١-عدم وجود قيم مفقودة .

٢-وجود قيمة مفقودة أو اكثر .

ب-في حالة فقدان قيمة واحدة أو اكثر ستتبع اربع طرق لتقدير القيم المفقودة هي Haseman and Gaylar, Rubin , Harry , Yates .

ج-مقارنة النتائج عن طريق قيم Cal,F , Mst ,Mse كما يلي :-

١- اسلوب تحليل الإنحدار باستخدام الترميز التاثيري وبين الطرق الكلاسيكية المتبعة للوصول الى جدول تحليل التباين عند عدم وجود قيم مفقودة .

٢- المقارنة بين النتائج المستحصلة عن طريق استخدام الترميز التاثيري في تحليل الإنحدار والطرق الاربع المشار اليها في (ب) عند وجود قيم مفقودة .

2

الفصل الثاني

الجانب النظري

الفصل الثاني

الجانب النظري

٢-١ التعاريف الأساسية في التجارب المصممة :

٢-١-١ التجربة Experiment :-

وهي الخطط التي ترسم بوسائل علمية لتشكل أساس جيد ومأمون لكي يـتم الحصـول على معلومات جديدة بين المتغيرات لرفض أو تأكيد فرضيات سابقة أو استنتاج قواعد وأسس جديدة (المشهداني والمشهداني ،٢٠٠٢) والتجارب أنواع منها البسيطة التي تهتم بدراسة تأثير عامل واحد أو يكون الاهتمام بدراسة تأثير أكثر من عامل وتسمى تجارب عاملية.

٢-١-٢ الوحدة التجريبية Experimental unit :-

أصغر قسم من مواد التجربة توزع عليها المعاملات المطلوبة لدراسة تأثيرها عشوائيا بحيث تطبق كل معاملة على وحدة تجريبية أو أكثر (دأود وعبد الياس ، ١٩٩٠) وقد تمثل التجربة مساحة من الأرض أو هي إنسان أو حيوان أو نبات وذلك يتوقف على نوع التجربة .

٢-١-٣ المعاملات (المعالجات) Treatments:-

وهي مجموعة من الظروف التجريبية التي توضع تحت سيطرة الباحث وتوزع عليها الوحدات التجريبية أو توزع على الوحدات التجريبية (المشهداني والمشهداني ،٢٠٠٢) وأن هذه الظروف التجريبية يمكن تقدير تأثيرها على صفة محددة لمواد التجربة ويمكن تمثيلها بمستويات عامل واحد أو أكثر وربما تكون المعاملات تحت الدراسة تمثل بمعاملات كمية أو وصفية .

٢-١-٤ الخطأ التجريبي Experimental Error:-

هو وصف الاختلاف الناتج من وحدات تجريبية متجانسة ومستقلة والاختلافات الاصلية للخطأ التجريبي (Kuehl , 2000) ومن ضمنها الإختلاف الطبيعي بين وحدات التجربة أو الاختلاف في قياس تطبيق المعاملة أو كفاءة الأشخاص الذين يقومون بتطبيق المعاملة أو الخطأ في القياسات أو عدم الحصول على وحدات متجانسة أو التفاعل بين المعاملات ووحدات التجربة وبقية العوامل ذات التأثير للصفات والاختلافات الوراثية أو التداخل بين التراكيب الوراثية والظروف البيئية التي يصعب السيطرة عليها.

٢-٢ تحليل التباين Analysis of Variance :-

هو الأسلوب الرياضي الذي يحدد فيما لو كان هناك اختلاف بين مجتمع إحصائي واحد أو أكثر والمقارنة بينهما والذي تتم بموجبه تجزئة مجموع المربعات الكلي لمجموعة من البيانات الى مصادره المختلفة (Keller and Warrock ,2003) والمسؤولة عن وجوده وتلخص النتائج في جدول بعد الإنتهاء من التحليل الذي يسمى جدول تحليل التباين ANOVA Table

٢-٢-١ الفروض الأساسية في تحليل التباين :

The Basic Assumptions of ANOVA

على الباحث أن يتأكد من البيانات المتوفرة لديه قبل ان يقوم بتحليلها من حيث استيفائها للشروط اللازمة لإجراء تحليل التباين حيث ان الخلل في هذه الشروط يؤدي الى عدم دقة النتائج وبالتالي ارتفاع مستوى المعنوية للإختبارات تلقائيا ومن ثم الوصول الى قرارات خاطئة ومن الشروط الواجب توفرها في تحليل التباين كما يبينها (الرأوي وخلف الله 1980) .

١. التأثيرات الأساسية تجميعية Additivity of the Main Effect

٢. التوزيع العشوائي المستقل والطبيعي للخطأ التجريبي

Randomly , Independently and Normally Distribution of the Error Term

٣. تجانس تباينات العينات المختلفة Homogeneity of Variance

٤. الاستقلالية بين المتوسطات والتباينات Independence of Means and Variance

٢-٢-٢ اسس تصميم التجارب Principle of Experimental Design :-

Kuehl(2000)

أوضح Fisher اسس تصميم التجارب الحديثة وقواعدها في مجال الزراعة وذلك عام 1925 في كتاب (Statistical Principles for Research Worker) وكتاب آخر لنفس المؤلف عام 1935 The Design of Expriments وقد عارض الباحثون الزراعيون في أول الأمر استعمال التجارب العشوائية التي أشار أليها (Fisher) وفضلوا عليها التجارب المنتظمة لصعوبة الأولى نسبة الى الأخيرة لكن سرعان ما حلت التجارب العشوائية محل التجارب المنتظمة و اصبحت تستعمل في التجارب الحقلية والمعملية وهناك ثلاثة متطلبات تعتمد عليها أسس تصميم التجارب وهي :

1. التوزيع العشوائي Randomization :

يعني اجراء التوزيع للمعاملات على الوحدات التجريبية بشكل عشوائي بعيداً عن التحيز وحسب نوع التصميم المستخدم في أية تجربة ويمكن باتباع هذا الاسلوب أن نتجنب الخطأ المنتظم ونضمن تقدير دقة الخطأ التجريبي وما يتسبب عنه في زيادة في كفاءة التجربة. داؤد وعبد الياس (١٩٩٠) .

2. التكرار Replication :

يقصد بالتكرار تمثيل المعاملة الواحدة في اكثر من وحدة تجريبية للحصول على فكرة صحيحة عن تأثير المعاملة وامكانية تقدير الخطأ التجريبي وزيادة كفاءة التجربة ودقتها ولزيادة تعميم نتائج التجربة وتقليل قيمة الخطأ التجريبي .

3 . التحكم في الوحدات التجريبية Local Control :

يتحدد عدد الوحدات التجريبية بعدد تكرار المشاهدات لكل وحدة تجريبية و ايضاً بعدد التكرارات في التجربة مما يحقق زيادة في دقة التجربة وذلك لزيادة عدد مرات ظهور المعاملة في التجربة بزيادة عدد مكررات التجربة أو بتكرار تسجيل مشاهدات على نفس الوحدة التجريبية التي عوملت بمعاملة معينة .

٢-٣-١ التصميم العشوائي الكامل (CRD) :

يعد من ابسط أنواع التصاميم وطريقة التحليل الإحصائي للبيانات سهلة حتى في حالـــة فقدان بعض المشاهدات الا انه يعاب عليه بعدم استخدامه الا في حالة تجانس وحدات التجربة وقيمة الخطأ التجريبي عالية مقارنة ببقية التصاميم مما ينتج عنه قلة في كفاءة التجربـــة دأود وعبد الياس (1990) والنموذج الرياضي لهذا التصميم

$$Y_{ij} = \mu + T_i + \in ij \qquad .....(1-2)$$

$$i = 1,2...t$$
$$j = 1,2...r$$

Y_{ij} : قيمة أو نتيجة القطعة (الوحدة) التجريبية الواقعة تحت تأثير المعاملة i والتي ترتيبها (j)

أو تسمى الإستجابة (response)

u : تأثير المتوسط العام $\hat{\mu} = \frac{Y..}{tr} = \bar{y}..$

Ti : تاثير المعالجة $\hat{T}i = \frac{Yi. - r\hat{\mu}}{r} = \frac{Yi.}{r} - \frac{Y..}{tr} = \bar{y}i. - \bar{y}..$

(i)

$\in ij$: الخطأ العشوائي للوحدة التجريبيةj التي اخذت المعاملة i

$$\in_{ij} = Y_{ij} - \hat{\mu} - \hat{T}i = y_{ij} - \bar{y}_i.$$

وجدول تحليل التباين لهذا التصميم

جدول رقم (1-2)

تحليل التباين للتصميم العشوائي الكامل

Source	*D.f*	*S.S*	*MS*	*F*
Treatments	*t-1*	$SS_t = \sum_{i=1}^{t} r_j (\bar{y}_i. - \bar{y}..)^2$	MS_t	$\frac{MSt}{MSe}$
Error	*t(r-1)*	$SS_e = \sum_{i=1}^{t} \sum_{j=1}^{r} (y_{ij} - \bar{y}_i.)^2$	MS_e	
Total	*rt-1*	$SS_T = \sum_{i=1}^{t} \sum_{j=1}^{r} (Y_{ij} - \bar{y}..)^2$		

ان هذا الجدول تم تنظيمه بهذه الطريقة للمقارنة مع جدول تحليل التباين للإنحـــدار بطريقـــة المتغيرات الوهمية الذي سيتم استعماله لاحقاً .

٢-٣-٢ تصميم القطاعات العشوائية الكاملة (R. C. B. D) :

وهو التصميم الذي تجمع فيه الوحدات التجريبية في مجاميع وتقسم الى قطاعات بحيـــث تكون هذه الوحدات التجريبية متجانسة داخل كل قطاع وتوزع المعاملات توزيعـــاً عشـــوائياً ومستقلاً داخل كل قطاع ويحتوي القطاع على جميع المعاملات . الراوي وخلف الله (1980) والنموذج الرياضي لهذا التصميم :

$$Y_{ij} = \mu + T_i + \rho_j + \in_{ij} \qquad ..... \quad (2-2)$$

$$i = 1,2...t \ , \ j = 1,2...r$$

Y_{ij} : قيمة المشاهدات للوحدة التجريبية التي أخذت المعاملة (i) الموجودة في القطاع (j)

u: قيمة الوسط العام

Ti:قيمة التأثير الحقيقي للمعاملة (i)

ρ_j:قيمة التأثير الحقيقي للقطاع (j)

$\in_{ij}$: قيمة التأثير الحقيقي للخطأ التجريبي الخاص بالمشاهدة ضمن المعاملةi والقطاع(j)

جدول (٢-٢)

جدول تحليل التباين لتصميم القطاعات العشوائية الكاملة

Source	$\partial.F$	s.s	M.s	F
Block	*r-1*	$SSr = t\sum_j (\bar{y}._j - \bar{y}..)^2$		
Treatments	*t-1*	$SS_t = r\sum_{i=1} (\bar{y}_i. - \bar{y}..)^2$	$Ms_t = \frac{ss_t}{t-1}$	$\frac{Ms_t}{Ms_e}$
Error	*(r-1)(t-1)*	$SS_e = \sum_{j=1}^{r}\sum_{i=1}^{t}(Y_{ij} - \bar{y}._j - \bar{y}_i. + \bar{y}..)^2$	$Ms_e = \frac{ss_e}{(t-1)(r-1)}$	
Total	*rt-1*	$SS_T = \sum_{j=1}^{r}\sum_{i=1}^{t}(y_{ij} - \bar{y}..)^2$		

ان هذا الجدول تم تنظيمه بهذه الطريقة مقارنة بجدول تحليل التباين للإنحدار باستخدام طريقة المتغيرات الوهمية الذي سيتم استعماله لاحقاً

٢-٣-٣ تصميم المربع اللاتيني (LSD) :

وهو التصميم الذي تجمع فيه الوحدات التجريبية باتجاهين متعامدين أحدهما يسمى الصفوف (Rows) والأخر الأعمدة (Columns) وتظهر كل معاملة مرة واحدة في كل صف أو عمود ولهذا يكون الخطأ صغيراً مما يؤدي إلى دقة هذا التصميم وكفاءته (هيكس شارلز ١٩٨٤) .

النموذج الرياضي لهذا التصميم هو :

$$Y_{ij(k)} = \mu + R_i + C_j + T_k + \in_{ij(k)} \qquad ...(3.2)$$

$$i = 1,2,\cdots,r$$

$$j = 1,2,\cdots,c$$

$$k = 1,2,\cdots,t$$

$$r = c = t$$

μ : المتوسط العام

Ri: التأثير الحقيقي للصف I

Cj: التأثير الحقيقي للعمود j

Tk: التأثير الحقيقي للمعاملة k

$\in_{ij}(k)$: القيمة الحقيقية للخطأ التجريبي للوحدة التجريبية ضمن المعاملة k

جدول (3-2) تحليل التباين لتصميم المربع اللاتيني

Source	$\partial.f$	$S.S$	Ms	F
Rows	$r-1$	$SS_r = r\sum_i (\bar{y}_i. - \bar{y}...)^2$	MS_R	
Columns	$r-1$	$SS_c = r\sum_j (\bar{y}._j - \bar{y}...)^2$	MS_c	
Treatment	$r-1$	$SS_t = r\sum_k (\bar{y}_k - \bar{y}...)^2$	MS_t	$\frac{MSt}{MSe}$
Error	$(r-1)(r-2)$	$SS_e = SS_T - SS_r - SS_C - SS_t$	MS_e	
Total	r^2-1	$SS_T = \sum_i \sum_j (y_{ij} - \bar{y}...)^2$		

ان هذا الجدول تم تنظيمه بهذه الطريقة مقارنة مع جدول تحليل التباين للإنحدار باستخدام طريقة المتغيرات الوهمية الذي سيتم استعماله لاحقا

٢-٤ القيم المفقودة Missing Values :

يحصل فقدان للقيم في أي تجربة عندما لا يمكن السيطرة بشكل كامل على التجربة من مؤثرات مختلفة ففي بعض التجارب المختلفة يحصل فقدان لبعض المشاهدات نتيجة لتلف إحدى القطع التجريبية بسبب العوارض كالأمطار والعواصف والثلوج أو نتيجة مرض الحيوان أو النبات أو عندما يموت الحيوان أو تقضي الحشرات على نباتات معينة من التجربة أو عندما تكون هناك خسارة في عينات الحصاد أو في تجارب اخرى في المختبر كأن تكسر أواني التجربة وقد يحصل فقدان البيانات في مصنع ما نتيجة للتوقفات وهكذا فإن فقدان البيانات يحصل في شتى مجالات الحياة ولأسباب مختلفة .

ففي التصميم العشوائي الكامل نقوم بتحليل البيانات في حالة فقدان بعض المشاهدات على أساس عدم تسأوي التكرار وبذلك يمكن التغلب على هذه المشكلة بقدر معين أما في تصميم القطاعات العشوائي الكاملة وتصميم المربع اللاتيني فيتم إيجاد قيم تقديريه للقيم المفقودة للمحافظة على اتزان التصميم ودقة النتائج ويتم ذلك من خلال اختيار افضل قيمة تقديريه يمكن الحصول عليها وذلك لجعل الفرق بين مجموع مربعات الخطأ التجريبي المستخرج بعد تقدير القيمة المفقودة والخطأ التجريبي في حالة عدم وجود قيمة مفقودة (قيمة حقيقية) اصغر ما يمكن.

٢-٥ طرائق تقدير القيم المفقودة :

Methods of Estimating the Missing Values

سنتناول في هذا الموضوع عرض بعض طرائق تقدير القيم المفقودة في تصميم القطاعات العشوائية الكاملة (RCBD) والمربع اللاتيني (LSD)

٢-٥-١ طريقة يتز Yates Method:

عالج Allan and Wishart في عام 1930 مشكلة تقدير القيم المفقودة في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني ووضعا صيغة مناسبة لمعالجة قيمة مفقودة واحدة وذلك بجعل مجموع مربعات الخطأ اقل ما يمكن .

وتمكن Yates في عام 1933 من تطوير هذه الطريقة وذلك بتقدير عدة قيم مفقودة في تصميم القطاعات العشوائية الكاملة وتصميم المربع اللاتيني بتطبيق اسلوب التعويض المتتالي (Iterative). ويمكن اشتقاق صيغة تقديريه للقيمة المفقودة كالأتي :

أ–في تصميم القطاعات العشوائية الكاملة :

فلو فرضنا ان القيمة $\hat{Y}_{11}$ هي القيمة المفقودة وبالاعتماد على نموذج القطاعات العشوائية يمكن اشتقاق الصيغة كالأتي :

$$\in_{ij} = y_{11}^2 + \sum_{i=2}^{t}\sum_{j=2}^{r} Y_{ij}^2 - \frac{(Y_{1}. + Y_{11})^2 + \sum_{i=2}^{t} Y_{1.}^2}{r} - \frac{(Y_{.1} + Y_{11})^2 + \sum_{J=2}^{r} Y_{.j}^2}{t} + \frac{(Y_{..} + Y_{11})^2}{tr}$$

$$\frac{\partial e_{ij}}{\partial y_{11}} = 2y_{11} - \frac{2(Y_{1}. + Y_{11})}{r} - \frac{2(Y_{.1} + Y_{11})}{t} + \frac{2(Y_{..} + Y_{11})}{tr}$$

و بمسأواة المعادلة بصفر وحلها نحصل على

$$\hat{y}_{11} = \frac{tY_{1}. + rY_{.1} - Y_{..}}{(r-1)(t-1)} \qquad \ldots(4\text{-}2)$$

وتكون $\hat{y}_{11}$ التي تم استخراجها في المعادلة رقم (٢–٤) هي القيمة التقديرية للقيمة المفقودة في تصميم القطاعات العشوائية الكاملة باستخدام طريقة Yates

حيث ان $\hat{y}_{11}$: القيمة التقديرية للقيمة المفقودة

$Y_{i}.$: مجموع المعاملة التي تحتوي على القيمة المفقودة.

$Y_{.j}$: مجموع القطاع الذي يحتوي على القيمة المفقودة.

$Y_{..}$: المجموع العام

r : عدد القطاعات

t : عدد المعاملات

ب – في تصميم المربع اللاتيني (LSD):

نفرض ان المشاهدة Y_{111} هي المفقودة التي تقع في العمود والصف الأول والمعاملة الأولى مفقودة في تصميم مربع لاتيني r × r فيكون مجموع مربعات الخطأ كالآتي :

$$\sum\sum\sum e_{ij(k)}^2 = y_{111}^2 + \sum_{i=2}\sum_{j=2}\sum_{k=2} y_{ij(k)}^2 - \frac{1}{r}\Big[\,(Y_{1..} + y_{111})^2 + \sum_{i=2} Y_{i..}^2 + (Y_{.1.} + y_{111})^2$$

$$+ \sum_{j=2} Y_{.1.}^2 + (Y_{..1} + y_{111})^2 + \sum_{k=2} Y_{..k}^2\,\Big] + \frac{2(Y_{...} + y_{111})^2}{r^2}$$