الفرضية المعلوماتية مقابل الفرضية الصفرية الإحصائية: دراسة الإحصاء الكلاسيكي مقابل الإحصاء البايزى

أ.د. محمد إبراهيم محمد محمد
أستاذ علم النفس التربوي
كلية التربية – جامعة المنيا

الملخص:

يهدف البحث الى دراسة الفرضية المعلوماتية مقابل الفرضية الصفرية الإحصائية حيث التعريف بالفرضية المعلوماتية ، وأسباب التوجه نحو استخدامها ، والعوامل المؤثرة على انخفاض انتشارها مقارنة بالفرضية الصفرية الإحصائية ، تم استخدام المراجعة المنهجية للأدبيات العلمية التى تناولت تلك القضية وتكونت عينة البحث من (92) بحثا ودراسة ،وقد توصلت النتائج إلى أن إجراء اختبار دلالة الفرضية الصفرية (NHST) يواجه انتقادا أكثر تكرارًا وشيوعًا لأنه يحتوي على العديد من القيود التي عندما يصبح الباحثون أكثر وعيًا بها، فمن المرجح أن يبتعدوا عن استخدامها لأن هذا النهج لا يعالج اهتماماتهم البحثية بشكل مباشر. تصبح الفرضية المعلوماتية، النهج بايزي لاختبار الفرضيات، خيارًا يلجأ إليه الباحث لمعالجة اهتماماته البحثية لأنه يوفر فوائد ومرونة أكبر وأقرب إلى معالجة اهتماماته وأهدافه البحثية. ومن ثم أصبح التقييم البايزي لفرضية عدم المساواة المقيدة بديلاً جذابًا لاختبار الفرضيات، ومواكبة استبدال NHST والتحرك نحوه. في حين أن انتقادات تقييم فرضية العدم التقليدية تتزايد بشكل مطرد، تحدد الفرضية المعلوماتية الفرضية المقصودة بشكل مباشر. في حين أنه من غير الممكن استنتاج NHST بقبول فرضية العدم، فإن الفرضية المعلوماتية توفر وسيلة لتنفيذ فرضيات متعددة وتحديد مستوى الأدلة على التأثير المقيد بالترتيب.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

The informational hypothesis versus the statistical null hypothesis (NHST): examining classical statistics versus Bayesian statistics

The research aims to study the informational hypothesis versus the statistical null hypothesis, where the informational hypothesis is defined, the reasons for the trend towards using it, and the factors affecting the decrease in its spread compared to the statistical null hypothesis. A systematic review of the scientific literature that dealt with this issue was used, and the research sample consisted of (92) research and studies. The results found that the Null Hypothesis Significance Testing (NHST) procedure faces a more frequent and common criticism because it has so many limitations that as researchers become more aware of them, they are likely to move away from using them because the approach does not directly address their research interests. Informatics hypothesis, the Bayesian approach to hypothesis testing, becomes an option for the researcher to address his or her research interests because it provides greater benefits and flexibility and is closer to addressing his or her research interests and goals. Bayesian evaluation of the bound inequality hypothesis has thus become an attractive alternative for hypothesis testing, keeping pace with and moving toward replacing NHST. While criticisms of traditional null hypothesis evaluation are steadily increasing, the informational hypothesis directly specifies the intended hypothesis. While it is not possible to conclude NHST by accepting the null hypothesis, the informational hypothesis provides a way to implement multiple hypotheses and determine the level of evidence for a rank-restricted effect.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

مقدمة البحث ومشكلته:

تسعى العلوم النفسية والتربوية إلى فهم الظواهر النفسية والتربوية وتفسيرها والتنبؤ بها، بالإضافة إلى إيجاد حلول مناسبة للمشاكل النفسية والتربوية . لتحقيق أهدافها تستخدم المؤسسة العلمية التحليل الإحصائي بالاشتراك مع شكل محدد من التفكير يعتمد على الاستنباط والاستقراء. ويعتبر هذا التفاعل بين الاستدلال الاستنباطي والاستقرائي هو الجوهر الأساس للاكتشاف الإحصائي في العلوم (Box, 1976). ومع ذلك، يعتقد بعض العلماء أن التوازن بين الاستقراء والاستنباط غير متوازن لصالح إجراءات الاستدلال الاستنتاجي (Orlitzky, 2012). بالإضافة إلى ذلك، يبدو أن الممارسة الإحصائية أصبحت عملية آلية يقوم بها الباحثون حيث يميلون إلى استخدام المعرفة الإحصائية بشكل أعمى لاتخاذ قرارات بشأن بياناتهم (Gigerenzer، 1998، 2004).

في عام 1925، قدم فيشر اختبار أهمية الفرضية الصفرية Null hypothesis statistical testing (NHST) لفصل النتائج المثيرة للاهتمام عن غيرها بشكل موضوعي. ويقصد بـ NHST هي طريقة تحليل البيانات الأكثر استخدامًا في معظم التخصصات العلمية. ننظر إلى الفرق بين المعالجتين اللتين نلاحظهما في التجربة ونسأل أنفسنا: "ما هو احتمال ملاحظة فرق بين المجموعات؟" يُشار إلى هذا الاحتمال باسم " القيمة p " أو "احتمال الأهمية". عندما يكون هذا الاحتمال صغيرا بما فيه الكفاية، نحن واثقون من أن احتمال عدم وجود فرق بين المعالجات صغير جدا، وبالتالي، نستنتج أن التجربة تدعم الفرضية البديلة (أي فرضية العمل التي حفزت الدراسة). عندما يكون الاحتمال أكبر، يكون لدينا القليل من الأدلة لدعم الفرضية البديلة، على الرغم من أنها قد تظل صحيحة.

Mudholkar, Chaubey,2019, Goodman ,1999,) (Fisher,1925

في اختبار الفرضيات الإحصائية، يمكن أن يحدث نوعان من الأخطاء: الإيجابيات الكاذبة (أي الرفض غير الصحيح للفرضية الصفرية) والسلبيات الكاذبة (أي الفشل في رفض الفرضية الصفرية الخاطئة). يستخدم نهج NHST قيمة قطع تعسفية (عادة 0.05) للتحكم في المعدل الإيجابي الكاذب. يتم وصف النتائج ذات القيم p الأصغر من قيمة القطع بأنها "ذات دلالة إحصائية" أو "إيجابية"، في حين توصف النتائج ذات القيم p التي تساوي أو أكبر من القطع بأنها "غير هامة" أو "سلبية". (Lu & Belitskaya-Levy,2015)

لاحظ الباحث من خلال الاطلاع على الأدبيات أن أدب علم النفس مليء بقيم الدلالة الإحصائية. في كل مقالة بحثية منشورة تقريبًا في علم النفس التربوى

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

والتربية، يتم دعم الادعاءات الموضوعية والدلالة الإحصائية، مما يشكل هيمنة للدلالة الإحصائية على الأدب النفسي والتربوي، حيث إن هيمنة الإطار الإحصائي لقيم الدلالة يأخذ منحى توجيهى حيث إن وجودها يبدو توجيهيًا تقريبًا ، فالباحث يسارع نحو الدلالة الإحصائية ومجرد الحصول عليها يجعله يستقر ويهدأ ويطمئن لدقة نتائجه وزيادة احتمال نشر ورقته البحثية .

وعلى الرغم من شعبيتها المستمرة، كانت القيمة p موضوعًا لعدد متزايد من الانتقادات. نذكر بإيجاز مسألتين فقط لهما أهمية خاصة في أبحاث التربية. أولاً، نظرًا لأن نهج القيمة p التقليدي يهتم فقط بدحض الفرضية الصفرية، فلا توجد طريقة لتقييم ما إذا كانت البيانات تفضل الفرضية الصفرية مقارنة بالفرضية البديلة. وحتى " كبيرة " لا توفر القيمة p غير المهمة دليلاً على فرضية العدم. وبالتالي، فإن فحص التكافؤ الإحصائي هو أمر بعيد عن متناول طرق اختبار القيمة الاحتمالية التقليدية ، ثانيًا ما لم يتم إجراء تحليل مسبق للقوة الإحصائية، فلا يوجد مؤشر واضح إذا كانت مجموعة البيانات حساسة بما يكفي لاكتشاف التأثير الحقيقي عند استخدام القيم P ومستويات ألفا ضمن المدرسة الكلاسيكية للإحصاءات التكرارية، وتستخدم للتحكم في معدلات الخطأ على المدى الطويل. (Wagenmakers, et. el, 2018)

ومن بين أبرز مشاكل القيمة الاحتمالية يمكننا تسليط الضوء على أنها تستخدم بطريقة الكل أو لا شيء لاتخاذ قرار بشأن الفرضيات الإحصائية أو أنه تم فهمها على أنها احتمالية الفرضية الصفرية صحيحة بشكل عام، يبدو أن هذه المشكلات المتعلقة بالقيمة الاحتمالية تنشأ في السياق التعليمي. على سبيل المثال، في دراسة أجراها (Haller & Krauss (2002 تبين أن طلاب وأساتذة الجامعات لديهم مفاهيم خاطئة مماثلة حول معنى ومنطق القيمة p. وهذا يوضح سوء الفهم للدلالة الإحصائية ومدلولاتها لدى المهتمين بالبحث العلمى حيث يتسارعون نحو الدلالة الإحصائية باعتبارها محك ثقة فى نتائج أبحاثهم .

وكما ذكر (Krzywinski & Altman 2013) فإن وجهة النظر الإحصائية هذه "لا تخبرنا ما إذا كنا على حق أم لا. إنه يخبرنا باحتمالات الخطأ". التوصية العامة هي استخدام القيمة p بشكل أكثر منطقية لتجنب التفسيرات الخاطئة، وتوفير أحجام التأثير، وتقديرات القوة وفترات الثقة .

وبالتالى توجهت أنظار العلماء والباحثين نحو محاولة إيجاد حلول للانتقادات المتزايدة التى تواجه الدلالة الإحصائية وتمثلت بعض الحلول فى ايجاد أساليب تكميلية بعد الدلالة الإحصائية مثل حجم الأثر والأساليب البعدية وأساليب قبلية مثل : القوة الإحصائية والتحديد الدقيق لحجم العينة والمعالجات الإحصائية ، وفى الجانب الآخر ظهرت بدائل للدلالة الإحصائية والإحصاء التكراري.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

ردًا على تلك الانتقادات المتزايدة للدلالة الإحصائية والإحصاء التكرارى فقد أصدرت المجلات الرائدة مثل Psychology Science مقالات افتتاحية تحذر من الاستخدام الحصري وغير النقدي للقيم الاحتمالية (Lindsay, 2015)؛ وقد تم تقديم تحذيرات مماثلة في النشرة النفسية ومراجعة الإرشادات الإحصائية للمؤلفين؛ أخيرًا، حظرت مجلة علم النفس الاجتماعي الأساسي والتطبيقي القيم p تمامًا (Trafimow & Marks, 2015).

وتعد أحد البدائل المقترحة للإحصاءات الكلاسيكية المستندة إلى NHST هو استخدام الاستدلال البايزي حيث إنه في سياق اختبارات الفرضيات الصفرية والبديلة، يقيس عامل بايز مقدار الأدلة التي تؤيد فرضية واحدة ضد فرضية أخرى. على الرغم من تطور الفرضية المعلوماتية منذ أكثر من ثلاثة عقود، إلا أن تفصيل عرض هذا الموضوع يقتصر بشكل رئيس على بيان التمثيل الفني والإحصائي في حين يتم التركيز على استخدام البرمجيات لإجراء الاختبار وإظهار الاهتمام العملي لإثبات التطبيق بشكل منهجي. على الرغم من تطور الفرضية المعلوماتية منذ أكثر من ثلاثة عقود، إلا أن تفصيل عرض هذا الموضوع يقتصر بشكل رئيس على بيان التمثيل الفني والإحصائي في حين يتم التركيز على استخدام البرمجيات لإجراء الاختبار وإظهار الاهتمام العملي لإثبات التطبيق بشكل منهجي.

لذا لفت انتباه الباحث الانتقادات العديدة التى وجهت للدلالة الإحصائية فى الأدبيات البحثية وطرح الإحصاء البايزى كبديل لحل المشكلات التى وجهت للإحصاء التكرارى التقليدى إلا أنه مازالت الهيمنة فى البحث العلمى والتربوى للدلالة الإحصائية والإحصاء التكرارى والفرضية الصفرية مازالت لها السيادة والسيطرة لذا وجد الباحث نفسه أمام العديد من الأسئلة التى يريد التوصل إلى إجابة من خلال المراجعة المنهجية للأدبيات التى تناولت تلك القضايا وتمثلت تلك الاسئلة فيما يلي:

1- ما الفرضية المعلوماتية مقابل الفرضية الصفرية الإحصائية ؟
2- لماذا الفرضية المعلوماتية ؟
3- لماذا تنتشر الفرضية الصفرية الإحصائية دون الفرضية المعلوماتية البيزية ؟

الأهـــداف :

يهدف البحث الحالى إلى :

- التعرف على الفرضية المعلوماتية مقابل الفرضية البيزية .
- تفسير أهمية الفرضية المعلوماتية مقابل الفرضية الصفرية .
- تفسير أسباب عدم شيوع الفرضية المعلوماتية مقابل الفرضية الصفرية.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

الأهمية :

تتضح أهمية البحث الحالى من خلال استعراضه لمفاهيم جديدة نسبيا على المجتمع العلمى التربوى والنفسي فى البيئة المصرية حيث يتناول الفرضية المعلوماتية مقابل الفرضية الصفرية، مما يساعد الباحث التربوى والنفسي العربي والمصري فى الإلمام بتلك التوجهات الحديثة فى المعالجات الإحصائية ويدفعه نحو المزيد من الدراسات والبحوث حول تلك القضايا المهمة .

مصطلحات البحث :

الدلالة الإحصائية :

القيمة p هي احتمال مشروط. على وجه التحديد، هو احتمال ملاحظة بيانات العينة، أو بعض الاحتمالات الأكثر تطرفًا، إذا كانت الفرضية الصفرية صحيحة. في المصطلحات المتكررة، تمثل القيمة p عدد المرات التي نلاحظ فيها عينة إحصائية، أو أكثر تطرفًا، في حالة كانت الفرضية الصفرية صحيحة في عدد السكان بعد أن كررنا تجربة واحدة عدة مرات في نفس الظروف. ونتيجة لذلك، كلما انخفضت القيمة p، زاد احتمال ظهور البيانات المرصودة في حالة صحة الفرضية الصفرية. وهذا هو السبب وراء اعتبار NHST الكلاسيكي لا يقدم معلومات حول حقيقة الفرضية الإحصائية. (García & Puga,2018)

عـامل بايز :

تعد أحد البدائل المقترحة للإحصاءات الكلاسيكية المستندة إلى NHST هو استخدام الاستدلال البايزي تجمع إحصائية بايزي بين البيانات المرصودة والمعلومات السابقة حول الظواهر لإجراء استنتاجات باستخدام نظرية بايز. ويعتبر نهجًا منطقيًا للعثور على الأدلة الداعمة للاستدلالات الإحصائية على عكس تقريب القيمة p، تركز إحصائيات بايزي على احتمالية الفرضية الإحصائية في ضوء بيانات العينة، داخل إطار الإحصاءات البايزية، حظيت عوامل بايز بالكثير من الاهتمام كبدائل لإجراء NHST يمثل عامل بايز مدى احتمالية فرضية العدم في ضوء البيانات المرصودة مقارنة بالفرضية البديلة. وبعبارة أخرى، "إنه العامل الذي يتم من خلاله تغيير الاحتمالات السابقة إلى الاحتمالات اللاحقة" لحدث ما (Bolstad, 2007)

الإجراءات :

منـهجية البحث :

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

اعتمد البحث المنهج الوصفى حيث تم البحث عن الاجابة عن اسئلة البحث من خلال الاطلاع على الأدبيات التى تناولت الفرضية المعلوماتية والفرضية الصفرية فى الإحصاء التكرارى والإحصاء البايزى وذلك من خلال المراجعة المنهجية للادبيات التى تناولت تلك القضية مع التركيز على ايجاد الاجابات للاسئلة التى طرحها الباحث من تلك الأدبيات وتجميعها معا .

عينة البحث :

تكونت عينة البحث من عدد (92) بحثًا ودراسة تناولت الفرضية المعلوماتية والفرضية الصفرية معا .

نتائج البحث :

الإجابة عن السؤال الأول : " ما الفرضية المعلوماتية مقابل الفرضية الصفرية الإحصائية ؟" تم الاطلاع على الأدبيات والبحوث الاجنبية فى التربية ومناهج البحث وأشارت إلى أن اختبار الفرضيات يعد إجراءً حاسمًا في الأبحاث الحديثة وينطبق في جميع المجالات والتخصصات تقريبًا، لا سيما عندما يكون اهتمام البحث هو اختبار احتمالية وجود اختلافات بين المجموعات فيما يتعلق بالمساواة وعدم المساواة وترتيب المجموعة، أصبح انتقاد نهج NHST أكثر تكرارًا وشيوعًا ؛لأنه يحتوي على العديد من القيود التي عندما يصبح الباحثون أكثر وعيًا بها، فمن المرجح أن يبتعدوا عن استخدامها ؛لأن هذا النهج لا يعالج اهتماماتهم البحثية بشكل مباشر. تصبح الفرضية المعلوماتية، النهج بايزي لاختبار الفرضيات، خيارًا يلجأ إليه الباحث لمعالجة اهتماماته البحثية ؛لأنه يوفر فوائد ومرونة أكبر أقرب إلى معالجة اهتماماته وأهدافه البحثية.(**Tan,2023**)

الفرضية المعلوماتية The informational hypothesis

تُعرف الفرضية المعلوماتية The informational hypothesis أيضًا باسم الفرضية المقيدة لعدم المساواة the inequality-constrained hypothesis .

ويحتوي على قيود لتحديد اختبار الفرضيات حول ترتيب المتوسطات ومعاملات الانحدار والمعلمات الإحصائية. تلك القيود تتمثل فى إدراج أحجام التأثير وقيود النطاق في اختيار قيود الفرضية المعلوماتية. يتضمن نوع القيود لصياغة فرضية معلومات أكبر من "<"، وأصغر من ">"، ويساوي "="، و"&"، وناقص "-"، بالإضافة إلى "+" هذا المكان بين المعلمات المحددة. (Hoijtink, Mulder & Van Lissa, et al. 2019; Van de Schoot, 2010).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

يحدد Hoijtink, Mulder & Van Lissa, et al (2019) مساحة الفرضيات المعلوماتية من خلال تحديد مواصفات العلاقات المتوقعة بين المعلمات (مثل المتوسطات) المصنفة ضمن أربع وحدات أساسية. تؤكد لبنة البناء الأولى على استخدام قيود المساواة والنظام بين المعلمات. الكتلة الثانية هي تحديد قيود المساواة والترتيب بين مجموعات المعلمات. تحدد الكتلة الثالثة استخدام حجم التأثير، وتسلط الكتلة الرابعة الضوء على قيد النطاق.

فرضية معلوماتية ومكملة وغير مقيدة The informational, complementary and unrestricted hypothesis. في حين أن NHST التقليدية لا تحدد أي تأثير للفرضية الصفرية والفرضية البديلة هي عدم تضمينها في الفرضية الصفرية، فإن الفرضية المعلوماتية تقدم مصطلحين فرضيين جديدين لم يتم اعتمادهما في النهج المتكرر. توضح الفرضية المعلوماتية قيود (عدم) المساواة لمعلمات النموذج . تتم مقارنة هذه الفرضيات إما بفرضية غير مقيدة أو فرضية مكملة.

الفرضية غير المقيدة The unrestricted hypothesis والتي يشار إليها أيضًا باسم الفرضية الشاملة encompassing hypothesis ، لا تضع أي قيود على النموذج الذي يمثل فرضية بديلة تغطي كل ترتيب لقيم المعلمات التي لا تتوافق مع الفرضية الأصلية. في سياق مقارنة متوسطات المجموعة، لا تفرض الفرضية غير المقيدة قيودًا على المتوسطات (Hoijtink, 2012).

الفرضية التكميلية The complement hypothesis هي المواصفات التكميلية لمجموعة محددة من الفرضيات المعلوماتية. إنها فرضية بديلة تغطي كل ترتيب لقيم المعلمات التي لا تتوافق مع مجموعة الفرضيات الأصلية (Böing-Messing, Van Assen, Hofman, et al, 2017)

عامل بايز والتعقيد والملاءمة Bayes Factor, Complexity, and Fit . يعد عامل بايز (BF) واحدًا من أقدم المؤشرات وأكثرها استخدامًا لتنفيذ اختبار الفرضية ضمن إطار بايزي. فهو يقارن القدرة التنبؤية لنموذجين متنافسين يتوافقان مع الفرضيات ويشير إلى درجة الأدلة في مجموعة البيانات بين الفرضية الصفرية والفرضية البديلة (Jeffreys, 1935). هناك اهتمام متزايد بعامل بايز لتقييم الفرضيات المقيدة كما يظهر من دراسات :

(e.g. Gu, Hoijtink, Mulder, et al, 2019; Gu, Hoijtink, Mulder, et al, 2020; Hoijtink, 2012; Hoijtink, Klugkist, & Boelen, 2008; Hoijtink, Mulder, Van Lissa & Gu, 2019; Kato & Hoijtink, 2006, Klugkist & Hoijtink, 2007; Klugkist, Laudy &

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

Hoijtink, 2005, Klugkist, Laudy, & Hoijtink, 2010, Laudy & Hoijtink, 2007; Mulder, Mulder, Hoijtink & Klugkist, 2010; Klugkist, Van de Schoot, Mulder, Hoijtink, et al, 2011; Wei, Yang, Rocha, et al, 2022)

في سياق الفرضية المعلوماتية ، فإن عامل بايز هو قياس كمي لمستوى ودرجة الأدلة في البيانات المجمعة لصالح الفرضية المعلوماتية.

الإجابة عن السؤال الثانى : "لماذا استخدام الفرضيات المعلوماتية ؟"

تم استعراض الأدبيات البحثية ،واتضح أن الأدبيات المتعلقة بالفرضيات المعلوماتية آخذة في الارتفاع وهذا الاتجاه المتزايد لا يخلو من سبب وجيه. هناك على الأقل ثلاث فوائد رئيسة لجأ إليها الباحثون لاختبار الفرضيات المعلوماتية .

أولاً: لدى العديد من الباحثين توقعات محددة حول نتائج أبحاثهم، والتي غالبًا ما يمكنها صياغة فرضياتهم بشكل لا لبس فيه. بالمقارنة مع NHST، فإن الفرضيات المعلوماتية يمكن أن تكون أكثر بكثير ربط توقعات الباحث بمجموعة من الفرضيات بسهولة أكبر باستخدام المتباينة، وتحديد الترتيب/الاتجاه، لتوضيح الفرضيات بشكل مباشر ومحدد. وتبين أن الفرضية المعلوماتية هي الطريقة الأنسب التي تسمح بتحديد البنية الشكلية للفرضيات لتتوافق مع أهداف البحث. تصبح فرضية الاختبار مباشرة تحت الفرضية المعلوماتية بدلاً من المواصفات غير المباشرة في NHST. على سبيل المثال، يمكن إجراء دراسة تهدف إلى اختبار ما إذا كان متوسط حجم التأثير في المجموعة الضابطة أكبر بثلاث مرات من متوسط مجموعة التجربة من خلال تحديد فرضية معلوماتية مرتبة لا يمكن تنفيذها في ظل NHST.

(e.g. Altinisik, Van Lissa, Gu, Hoitink, et al, 2021; Boehm, Steingroever & Wagenmakers, 2018; Hoijtink, 1998, 2000, 2001, 2012, 2013; Hoijtink, Klugkist, & Boelen, 2008; Hoijtink, Muder, Van Lissa, & Gu, 2008; Klugkist & Hoijtink, 2007; Klugkist, Laudy, & Hoijtink, 2005; Kuiper, Klugkist, & Hoijtink, 2010; Kuiper, Hoijtink & Silvapulle, 2012; Laudy & Hoijtink, 2007; Laudy, Boom, & Hoijtink, 2005; Mulder, Hoijtink, & Klugkist, 2010; Mulder, William, Gu, et al, 2021; Klugkist, Laudy & Hoijtink, 2005; Klugkist, Laudy, & Hoijtink, 2010; Rouder, Speckman, Sun, et al, 2009; Van de Schoot, 2010; Vanbrabant, 2020; Vanbrabant, Van de Schoot, &

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

Rosseel, 2015; Vanbrabant, Van Rossum, Van de Schoot, & Hoijtink, 2013; Wei, Yanf, Rocha, et al, 2022)

ثانيًا: ربما تكون الميزة الكبرى لاستخدام الفرضية المعلوماتية هي أن اختبار الفرضيات المتعددة ممكن دائمًا، حيث يعد تحديد سلسلة من الفرضيات سمة أساسية للفرضية المعلوماتية التي تسمح للباحث ليس فقط بوضع قائمة من الفرضيات ولكن أيضًا تحديد المستوى النسبي للأدلة لتحديد أفضل فرضية لمعالجة أسئلة البحث وأيضًا فيما يتعلق بالتصنيف والتقييم لمستوى الأهمية من خلال مجموعة من الفرضيات. Heck, Boehm, Böing-Messing, et al (2022)

أما الفائدة الثالثة من استخدام الفرضية المعلوماتية فلا تقتصر على مواصفات اختبار الفرضيات، بل تتعلق أيضًا بتصميم البحث. على سبيل المثال، أجرى مونين وسوير وماركيز (2008) تجارب من خلال مقارنة 27 مجموعة بأربعة 29 شرطًا باستخدام الفرضية المعلوماتية. استخدم Moerbeek (2019) فرضية ترتيبية للتدخل المدرسي للوقاية من التدخين مع أربع مجموعات علاجية. لا تساعد الفرضية المعلوماتية الباحث على تحديد الفرضية على وجه التحديد فحسب، بل تساعد أيضًا في تخطيط تصميم بحثه الذي يهدف إلى تقليل تكلفة البحث. والميزة هي أنه بالنظر إلى نفس حجم العينة، فإن قوة الفرضية المعلوماتية تكون أعلى. يمكن اكتساب القوة ومن المتوقع بطبيعتها أن يكون حجم العينة أصغر باستخدام الفرضية المعلوماتية، لا يقتصر اكتساب القدرة ومتطلبات حجم العينة الأصغر على مقارنة المتوسطات، بل ينطبق أيضًا على نماذج أخرى مثل نموذج المعادلة الهيكلية

(Van de Schoot & Strohmeier,2011, Vanbrabant, Van de Schoot & Roseel., 2015 , Van de Schoot & Strohmeier, 2011, Monin, Sawyer, & Marquez , 2008)

الإجابة عن السؤال الثالث :" لماذا تنتشر الفرضية الصفرية دون الفرضية المعلوماتية البيزية ؟"

اتضح للباحث من خلال الأدبيات: إنه قد يكون من الصعب فهم الشعبية المستمرة لقيم p مقارنة بالطرق البايزية. تم انتقاد مفهوم الاختبار الإحصائي لفرضية القيمة (NHST) مرارًا وتكرارًا على عدد من النقاط المهمة، وقد سعى عدد قليل من المنهجيين للدفاع عن هذه الممارسة. أحد الانتقادات هو أن القيم p غالبًا ما يتم إساءة تفسيرها على أنها احتمالات بايزي لاحقة، بحيث يكون من السهل جدًا الاعتقاد بأن $p < .05$ يستدعي رفض الفرضية الصفرية H0، وبالتالي يدعم قبول الفرضية البديلة. هذا التفسير لقيم p مغري ولكنه غير صحيح (Gigerenzer, Krauss, & Vitouch, 2004).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

حيث يشير (Wagenmakers, et. el, 2018) إلى أنه من منظور نفسي، قد يساعد عدد من الحجج في تفسير استمرار شعبية القيم الاحتمالية على الطرق البايزية ،ويقوم الباحث باستعراض تلك الحجج والبراهين الستة التى قدمها هنا حيث جاءت شاملة وكافية .

- **أولاً:** يمارس الباحثون ويبشرون بالمنهجية التي تعلموها ذات يوم بأنفسهم؛ يتطلب التغيير أن يستثمر الباحثون جهدًا جادًا لتعلم أساليب جديدة.

- **ثانيًا:** من خلال الانفصال عن المجموعة المهيمنة من ممارسي القيمة p، يختار الباحثون الابتعاد عن المجموعة الداخلية وتعريض أنفسهم للمخاطر المرتبطة بالاستبعاد الأكاديمي.

- **ثالثًا:** قد يعتقد الباحثون أن هناك أمانًا في تكرار الإجراءات الشائعة.

- **رابعًا:** يهتم العديد من علماء النفس في المقام الأول بمعالجة أسئلة البحث الموضوعية، وليس بالتفاصيل الدقيقة للمنهجية الإحصائية؛ إن عدم الاهتمام المنهجي هذا يغذي الرغبة في إجراءات بسيطة تعمل بشكل جيد بما يكفي لإقناع المراجعين. وبهذا المعنى، فإن تثبيت القيمة p الحالي يشبه الطقوس الإحصائية (Gigerenzer, Krauss, & Vitouch, 2004).

- **خامسًا:** يقدم إطار القيمة p، عند إساءة تفسيره، حلاً بسيطًا للتعامل مع عدم اليقين المتأصل في البيانات الصاخبة: عندما تكون p <.05، يتم رفض H0 وقبول H1 ؛ عندما p > .10، احتفظ بـ H0 . عند إساءة تطبيقها بهذه الطريقة، يبدو أن قيم P تجعل من السهل على الباحث استخلاص استنتاجات قوية حتى عندما تكون النتائج التجريبية صاخبة وغير مفيدة.

- **سادسا:** قد يشعر الباحثون أنه باستخدام أساليب غير قياسية (أي شيء آخر غير قيمة P) فإنهم يقللون من فرصهم في نشر أعمالهم أو فهمها من قبل زملائهم.

- **سابعا:** غالبا ما استوعب الباحثون المهتمون بالمنهجية تعليمهم الإحصائي إلى حد أنهم يجدون صعوبة في قبول أن الطريقة التي استخدموها طوال حياتهم قد تكون لها قيود خطيرة؛ عندما تتعارض المعلومات الجديدة مع العادات القديمة، يمكن تقليل التنافر المعرفي الناتج عن طريق استبعاد المعلومات الجديدة أو تجاهلها.

- **أخيرًا:** من الممكن أن يتفق الباحثون مع انتقادات القيمة p، لكنهم غير قادرين على اعتماد إجراءات استنتاجية بديلة (بايزي). والسبب وراء هذا العجز واضح

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

ومباشر: فكل حزم البرامج الإحصائية تقريبا تنتج قيم p بسهولة، في حين لا تستطيع الأساليب البايزية الاعتماد على نفس المستوى من الدعم. العديد من هذه الحجج تنطبق على الابتكارات الإحصائية بشكل عام، وليس فقط بالنسبة للقيمة الاحتمالية. (Sharpe, 2013)

الاستنتاجات :

الإحصاء البيزى يقدم نهجا بديلا؛ لأنه يسمح بالوصف الاحتمالي للمعلمات والفرضيات. كانت هناك العديد من المنشورات التي توضح بالتفصيل الاختلافات الفلسفية والعملية بين وجهتي النظر هاتين ، ولكن يكفي لأغراضنا أن نلاحظ أن الإطار البايزي فقط هو الذي يسمح لنا بتحديد مدى احتمالية وجود البيانات تحت الصفر مقارنة الفرضية الصفرية بالفرضية البديلة بشرط احتمال سابق. يمكن لعامل بايز، وهو تطبيق شائع لاختبار الفرضيات البايزية، أن يحدد مدى تفضيل البيانات لإحدى الفرضيتين من خلال النظر في الاحتمالات السابقة. من المهم ملاحظة أن إطار عمل بايزي يتضمن أيضًا تقدير المعلمة، والذي يمكنه معالجة حجم التأثير. في حين أن تقدير المعلمة البايزية يعد أداة قيمة، فإن اختبار الفرضيات عبر مقارنة النماذج البايزية يمكن أن يسهل التنبؤ النظري من خلال توفير مقياس للأدلة النسبية بين نموذجين ، وعادةً ما يكون نموذجًا صفريا وبديلاً.

(Gelman et. al.,2014 , McElreath,2015)

يعد تحديد توزيع مسبق للمعلمة في نموذج إحصائي أمرًا أساسيًا للاستدلال البايزي، ويخدم العديد من الأغراض مثل تقدير المعلمة المحسن . يمكن للتوزيع المسبق أن يحدد، أو على الأقل يقترب من المفهوم المثالي، للمعلومات السابقة حول معلمات النموذج قبل أخذ البيانات بعين الاعتبار. على عكس الأطر الاستدلالية الكلاسيكية، يمكن للاستدلال البايزي أن يتضمن مثل هذه المعرفة السابقة. على سبيل المثال: إذا كنا نتعامل مع معلمة حجم التأثير، مثل Cohen ' s d ، فقد نحكم مسبقًا على أن قيم d أقل من -1 أو أكبر 1، أقل احتمالًا بكثير من قيم d حول 0. وهذا افتراض عادل لأبحاث السلوك الحيوي، والتي تميل إلى تحقيق أحجام تأثير صغيرة إلى متوسطة (أي d = 0.2 إلى d = 0.5). إذا كانت المعلمة غير مقيدة، فإن استخدام توزيع كوشي المتمركز حول التأثير هو أسلوب شائع. عادةً ما يتمركز هذا التوزيع على الصفر بشكل افتراضي، ولكن يمكن أيضًا أن يتمركز في مكان آخر. يشبه توزيع كوشي التوزيع الطبيعي، ولكن له ذيول أكثر بدانة وكتلة مركزية أقل. التوزيعات العادية و t هي أيضًا خيارات شائعة للتوزيعات السابقة. (Chung,2012, Gelman et. al.,2017)

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

وبشكل عام هناك قوى نفسية واجتماعية قوية مؤثرة، مما يجعل من المستحيل تقريبًا تحدي المنهجية السائدة. ومع ذلك، يبدو أن صرح NHST يظهر عليه علامات الاضمحلال الدقيقة. يمكن القول: إن هذا يرجع إلى التجارب والمحن الأخيرة المعروفة مجتمعة باسم "أزمة الثقة" في البحث النفسي، وفي الواقع، في البحث التجريبي بشكل عام كما أشارت دراسات :

(Begley & Ellis, 2012; Button et al., 2013; John, Loewenstein, & Prelec, 2012; Nosek & Bar-Anan, 2012; Nosek, Spies, & Motyl, 2012; Pashler & Wagenmakers,2012 ; Simmons, Nelson, & Simonsohn, 2011).

إلى أن أزمة الثقة حفزت على إعادة التوجيه المنهجي بعيدًا عن الممارسة الحالية المتمثلة في القيمة P NHST. وقد شددت سلسلة من المقالات الحديثة على القيود المفروضة على القيم p واقترحت طرقًا بديلة للتحليل

(Cumming, 2008, 2014; Halsey, Curran-Everett, Vowler, & Drummond, 2015; Johnson, 2013; Kruschke, 2010a, 2011; Nuzzo, 2014; Simonsohn, 2015b)

ومن أجل تقليل اعتماد علماء النفس على القيم الاحتمالية، من الضروري تقديم بدائل ملموسة وعملية. أحد هذه البدائل هو الاستدلال من فترات الثقة حيث اقترحت الدراسات معالجات بعدية تكاملية مع الدلالة الإحصائية : مثل فترات الثقة .((Cumming, 2014; Grant, 1962

ورغم ذلك تم توجيه انتقادات لفترات الثقة حيث يرى الباحثون اثنين من القيود الرئيسة للإحصاءات الجديدة. القيد الأول هو أن فترات الثقة ليست بايزي، مما يعني أنها تتخلى عن الفوائد التي تأتي مع النهج بايزي (ترد قائمة بهذه الفوائد أدناه)؛ علاوة على ذلك، تشترك فترات الثقة في مصير القيم p، بمعنى أنها عرضة للمغالطات والتفسيرات الخاطئة (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016).

والقيد الثاني هو أن فترات الثقة تفترض وجود التأثير قيد النظر؛ وبعبارة أخرى، فإن استخدامها يعني أن كل مشكلة استدلال هي مشكلة تقدير المعلمة وليس اختبار الفرضيات. على الرغم من أننا نعتقد أن تقدير حجم التأثير مهم ويجب أن يحظى بالاهتمام، فإن مسألة الحجم ("ما هو حجم التأثير؟") لا تدخل حيز التنفيذ إلا بعد معالجة مسألة الحضور ("هل هناك تأثير؟") بشكل مقنع (Morey, Rouder, Verhagen, & Wagenmakers, 2014).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

ومن الحجج القوية والبراهين التى وجهت لانتقادات الأساليب التكميلية البعدية للدلالة الإحصائية إن حجم التأثير وفترات الثقة تأتى بعد استخدام الدلالة الإحصائية وتحقق الدلالة ،حيث أشار الرائد البايزي هارولد جيفريز في دراسته "نظرية الاحتمالية" إلى التمييز بشكل حاد بين التقدير والاختبار، ويناقش كل منهما في فصول منفصلة: "في مشاكل الفصلين الأخيرين كنا مهتمين بتقدير المعلمات في القانون، يتم إعطاء شكل القانون نفسه. نحن الآن معنيون بالسؤال الأكثر صعوبة: في أي ظروف تدعم الملاحظات تغيير شكل القانون نفسه؟ هذا السؤال منطقيًا يسبق تقدير المعلمات، حيث إن مشكلة التقدير تفترض أن المعلمات ذات صلة. (Jeffreys, 1961).

وقد اتفق (Simonsohn,2015)مع آراء هارولد جيفريز حيث أعرب عن نفس الشعور: فقط بمجرد أن نتجاوز التساؤل عما إذا كانت الظاهرة موجودة على الإطلاق ونقبلها على أنها صحيحة من الناحية النوعية، يمكننا أن نصبح مهتمين بتقدير حجمها بشكل أكثر دقة. قبل أن تصل خطوط البحث إلى الموقع المتميز المتمثل في تحديد ظاهرة مقبولة بشكل عام على أنها صحيحة من الناحية النوعية، يحتاج الباحثون إلى أدوات لمساعدتهم على التمييز بين تلك التي من المحتمل أن تصل إلى هناك وتلك التي ليس من المحتمل أن تصل إليها. نعتقد أنه من الخطأ فرض تقدير أو نهج اختبار في جميع المجالات؛ وبدلاً من ذلك، يعتمد أسلوب الاستدلال الأكثر إنتاجية على الأسئلة الموضوعية التي يرغب الباحثون في الإجابة عليها. كما هو موضح أدناه، فإن المشاكل المتعلقة بقيم الدلالة الإحصائية P ليست سببًا للتخلي عن اختبار الفرضيات - بل هي سبب للتخلي عن القيم p.

ويصل الباحث من خلال استعراض الأدبيات إلى أن التقييم البايزي لفرضية عدم المساواة المقيدة (الفرضية المعلوماتية) بديلاً جذابًا لاختبار الفرضيات، ومواكبة استبدال NHST والتحرك نحوه. في حين أن انتقادات تقييم فرضية العدم التقليدية تتزايد بشكل مطرد، تحدد الفرضية المعلوماتية الفرضية المقصودة بشكل مباشر. في حين أنه من غير الممكن استنتاج NHST بقبول فرضية العدم، فإن الفرضية المعلوماتية توفر وسيلة لتنفيذ فرضيات متعددة وتحديد مستوى الأدلة على التأثير المقيد بالترتيب.

ومن ثم يوصى البحث بمزيد من البحوث والدراسات العربية التربوية والنفسية والإحصائية التى تتناول التوجه نحو الفرضية المعلوماتية وفوائد استخدامها ونشر ثقافتها وكيفية تطبيقها وتوظيفها وتفسيرها مما يساعد الباحثين على التوجه نحو استخدامها كبديل للإحصاء التكراري.

المراجـع

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146

Altinisik, Y., Van Lissa, C. J., Hoijtink, H., Oldehinkel, A. J., & Kuiper, R. M. (2021). Evaluation of inequality constrained hypotheses using a generalized of the AIC. *Psychological Methods*, *26(5)*, 599-621. https://doi.org/10.1037/met0000406.

Altman, N., & Krzywinski, M. (2017a). Points of signifi cance: *P* values and the search for signifi cance. *Nature Methods*, *14*, 3-4. doi: 10.1038/nmeth.4120

Altman, N., & Krzywinski, M. (2017b). Points of signifi cance: Interpreting *P* values. *Nature Methods*, *14*, 213-214. doi: 10.1038/nmeth.4210

Begley, C. G., & Ellis, L. M. (2012). Raise standards for preclinical cancer research. *Nature*, *483*, 531–533.

Boehm, U., Steingroever, H. & Wagenmakers, E.-J. (2018). Using Bayesian regression to test hypotheses about relationships between parameters and covariates in cognitive models. *Behavior Research Methods*, *50*, 1248–1269. DOI: 10.3758/s13428-017-0940-4.

Böing-Messing, F., van Assen, M. A. L. M., Hofman, A. D., Hoijtink, H., & Mulder, J. (2017). Bayesian evaluation of constrained hypotheses on variances of multiple independent groups. *Psychological Methods*, *22*, 262–287. https://doi.org/10.1037/met0000116.

Bolstad, W. M. (2007). *Introduction to Bayesian statistics* (2nd ed.).Hoboken, NJ: Wiley.

Box, G. E. P. (1976). Science and statistics. Journal of the American Statistical Association, 71, 791-799.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munaf`o, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*, 1–12.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

Chung Y, Rabe-Hesketh S, Gelman A, Liu J, Dorie V(. 2012). Avoiding boundary estimates in linear mixed models through weakly informative priors. UC Berkeley Division of Biostatistics Working Paper Series. 2012; Working Paper 284.

Cumming, G. (2008). Replication and *p* intervals: *p* values predict the future only vaguely, but confidence intervals do much better. *Perspectives on Psychological Science*, *3*, 286–300.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*, 7–29.

Fisher RA.(1925). Statistical Methods for Research Workers. London:Oliver & Boyd

García, A. M. R. R., & Puga, J. L. (2018). Deciding on Null Hypotheses using P-values or Bayesian alternatives: A simulation study. Psicothema, 30(1), 110-115.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB.(2014). Bayesian data analysis, vol. 2. Boca Raton, FL: CRC press.

Gelman A, Simpson D, Betancourt M(2017). The prior can generally only be understood in the context of the likelihood. arXiv preprint arXiv:170807487. 2017.

Gigerenzer, G. (2004). Mindless statistics. The Journal of Socio-Economics, 33, 587-606. doi: 10.1016/j.socec.2004.09.033.

Gigerenzer, G., Krauss, S., & Vitouch, O. (2004). The null ritual:What you always wanted to know about significance testing but were afraid to ask. In Kaplan, D. (Ed.) *The sage handbook of quantitative methodology for the social sciences* (pp. 391–408). Thousand Oaks, CA: Sage.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

Goodman SN(1999). Toward evidence-based medical statistics. 1:The P value fallacy. Ann Intern Med. 1999; 130(12): 995-1004.http://dx.doi.org/10.7326/0003-4819-130-12-199906150-00008.

Grant, D. A. (1962). Testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychological Review*,*69*, 54–61.

Gu, X., Hoijtink, H., Mulder, J., Van Lissa, & Rosseel, Y. (2019). Bain: A program for Bayesian testing of order constrained hypotheses in structural equation models. *Journal of Statistical Computation and Simulation, 89(8)*, 1526-1553, DOI: 0.1080/00949655.2019.1590574.

Gu, X., Hoijtink, H., Mulder, J., Van Lissa, C. J., Van Zundert, C., Jones, J., & Waller, N. (2020). *Bain: Bayes factors for informative hypotheses* (R package version 0.2.4) [Computer software manual]. https://CRAN.R-project.org/package=bain.

Haller, H., & Krauss, S. (2002). Misinterpretations of signifi cance: A problem students share with their teachers? *Methods of Psychological Research Online*, *7*(1).

Halsey LG, Curran-Everett D, Vowler SL, Drummond GB. (2015)Thefickle P value generates irreproducible results. Nat Methods.2015; 12(3): 179-185. doi: http://dx.doi.org/10.1038/nmeth.3288

Halsey, L. G., Curran-Everett, D., Vowler, S. L., & Drummond, G. B. (2015). The fickle *P* value generates irreproducible results. *Nature Methods*, *12*, 179–185.

Heck, D. W., Boehm, U., Böing-Messing, F., Bürkner, P.-C., Derks, K., Dienes, Z., Fu, Q., Gu, X., Karimova, D., Kiers, H. A. L., Klugkist, I., Kuiper, R. M., Lee, M. D., Leenders, R., Leplaa, H. J., Linde, M., Ly, A., Meijerink-Bosman, M., Moerbeek, M., Mulder, J., Palfi, B., Schönbrodt, F. D.,

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

Tendeiro, J. N., van den Bergh, D., Van Lissa, C. J., van Ravenzwaaij, D., Vanpaemel, W., Wagenmakers, E.-J., Williams, D. R., Zondervan-Zwijnenburg, M., & Hoijtink, H. (2022). A Review of Applications of the Bayes Factor in Psychological Research. *Psychological Methods*. Advance online publication. http://dx.doi.org/10.1037/met0000454.

Held, L., & Ott, M. (2018). On p-values and Bayes factors. *Annual Review of Statistics and Its Application*, 5. doi: 10.1146/annurev-statistics- 031017-100307

Hoijtink, H. (2012). *Informative Hypotheses. Theory and Practice for Behavioral and Social Scientists*. Chapman & Hall/CRC.

Hoijtink, H., Gu, X., Mulder, J., & Rosseel, Y. (2019). Computing Bayes factors from data with missing values. *Psychological Methods, 24(2),* 253-268. http://dx.doi.org/10.1037/met0000187.

Hoijtink, H., Gu, X., Mulder, J., & Rosseel, Y. (2019). Computing Bayes factors from data with missing values. *Psychological Methods, 24(2),* 253-268. http://dx.doi.org/10.1037/met0000187.

Hoijtink, H., Klugkist, I., & Boelen, P. A. (Eds.). (2008). *Bayesian Evaluation of Informative Hypotheses*. New York, NY: Springer.

Hoijtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods, 24(5),* 539–556. http://dx.doi.org/10.1037/met0000201.

Hoijtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods, 24(5),* 539–556. http://dx.doi.org/10.1037/met0000201.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

Hoijtink, H., van Kooten, P., & Hulsker, K. (2016a). Why Bayesian psychologists should change the way they use the Bayes Factor. *Multivariate Behavioral Research, 51*, 2-10. doi: 10.1080/00273171.2014.969364

Hoijtink, H., van Kooten, P., & Hulsker, K. (2016b). Bayes factors have frequency properties-This should not be ignored: A rejoinder to Morey, Wagenmakers, and Rouder. *Multivariate Behavioral Research, 51*, 20- 22. doi: 10.1080/00273171.2015.1071705

Jarosz, A., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *Journal of Problem Solving, 7*, 2-9. doi: 10.7771/1932-6246.1167

Jeffreys, H. (1961). *Theory of probability*, 3rd edn. Oxford: Oxford University Press.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth–telling. *Psychological Science*, *23*, 524–532.

Johnson, V. E. (2013). Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 19313–19317.

Kato, B. S. & Hoijtink, H. (2006). A Bayesian approach to inequality constrained linear mixed models: Estimation and model selection. *Statistical Modeling, 6*, 231–249.

Klugkist, I., & Hoijtink, H. (2007). The Bayes factor for inequality and about equality constrained models. *Computational Statistics and Data Analysis, 51*, 6367–6379.

Klugkist, I., & Hoijtink, H. (2007). The Bayes factor for inequality and about equality constrained models.

Computational Statistics and Data Analysis, 51, 6367–6379.

Klugkist, I., Laudy, O. & Hoijtink, H. (2005). Inequality constrained analysis of variance: A Bayesian approach. *Psychological Methods, 10*, 477–493.

Klugkist, I., Laudy, O. & Hoijtink, H. (2005). Inequality constrained analysis of variance: A Bayesian approach. *Psychological Methods, 10*, 477–493.

Klugkist, I., Laudy, O. & Hoijtink, H. (2010). Bayesian evaluation of inequality and equality constrained hypotheses for contingency tables. *Psychological Methods, 15*, 281–299.

Klugkist, I., Laudy, O. & Hoijtink, H. (2010). Bayesian evaluation of inequality and equality constrained hypotheses for contingency tables. *Psychological Methods, 15*, 281–299.

Kruschke, J. K. (2010a). *Doing Bayesian data analysis: A tutorial introduction with R and BUGS Burlington*. MA: Academic Press.

Kruschke, J. K. (2010b). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, *14*, 293–300.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*, 299–312.

Krzywinski, M., & Altman, N. (2013). Importance of being uncertain.Nature Methods, 10, 809-810. doi: 10.1038/nmeth.2613

Kuiper R.M., Hoijtink H., Silvapulle M.J. (2012). Generalization of the Order-Restricted Information Criterion for Multivariate Normal Linear Models. *Journal*

of Statistical Planning and Inference, *142*, 2454–2463. DOI: 10.1016/j.jspi.2012.03.007.

Kuiper, R. M., & Hoijtink, H. (2010). Comparisons of means using confirmatory and exploratory approaches. *Psychological Methods, 15*, 69–86.

Kuiper, R. M., Altinisik, Y., & Van Lissa, C. J. (2021). gorica: Evaluation of inequality constrained hypotheses using GORICA. *R package version 0.1.2*, https://informative-hypotheses.sites.uu.nl/software/goric/.

Kuiper, R. M., Hoijtink, H., & Silvapulle, M. J. (2011). An Akaike-type information criterion for model selection under inequality constraints. *Biometrika, 98*, 495-501. http://dx.doi.org/10.1093/biomet/asr002.

Laudy, O., & Hoijtink, H. (2007). Bayesian methods for the analysis of inequality constrained contingency tables. *Statistical Methods in Medical Research, 16*, 123–138.

Laudy, O., Boom, J., & Hoijtink, H. (2005). Bayesian computational methods for inequality constrained latent class analysis. In A. Van der Ark & M. A. C. K. Sijtsma (Eds.), *New development in categorical data analysis for the social and behavioural sciences* (pp. 63–82). London, UK: Lawrence Erlbaum Associates, Ltd.

Lindsay, D. S. (2015). Replication in psychological science. *Psychological Science*, *26*, 1827–1832.

Lu, Y., & Belitskaya-Levy, I. (2015). The debate about p-values. Shanghai Archives of Psychiatry, 27(6), 381-385.

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis signifi cance testing. *Behavioral Research*, *43*, 679-690. doi: 10.3758/s13428-010-0049-5

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

McElreath R. (2015).Statistical rethinking: a Bayesian course with examples in R and Stan. Boca Raton, FL: CRC Press.

Morey, R. D., & Rouder, J. N. (2011). Bayes Factor approaches for testing interval null hypothesis. *Psychological Methods*, *16*, 406-419. doi: 10.1037/a0024377

Morey, R. D., Rouder, J. N., Verhagen, A. J., & Wagenmakers, E. J. (2014). Why hypothesis tests are essential for psychological science: A comment on Cumming. *Psychological Science*, *25*,1289–1290.

Morey, R. D., Wagenmakers, E., & Rouder, J. N. (2016). Calibrated Bayes factors should not be used: A reply to Hoijtink, van Kooten, and Hulsker. *Multivariate Behavioral Research*, *51*, 11-19. doi:10.1080/00273171.2015.1052710

Mudholkar GS, Chaubey YP(2009). On defining P-values. Stat Prob Letters. 2009; 79(18): 1963-1971. doi: http://dx.doi.org/10.1016/j.spl.2009.06.006

Mulder, J., Hoijtink, H., & Klugkist, I. (2010). Equality and inequality constrained multivariate linear models: Objective model selection using constrained posterior priors. *Journal of Statistical Planning and Inference, 140,* 887–906.

Mulder, J., Hoijtink, H., & Klugkist, I. (2010). Equality and inequality constrained multivariate linear models: Objective model selection using constrained posterior priors. *Journal of Statistical Planning and Inference, 140,* 887–906.

Mulder, J., Klugkist, I., Van de Schoot, R., Meeus, W., Selfhout, M., & Hoijtink, H. (2009). Informative hypotheses for repeated measurements: A Bayesian approach. *Journal of Mathematical Psychology, 53*, 530–546.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

Mulder, J., Williams, D. R., Gu, X., Andrew Tomarken, Boing-Messing, F., Olsson-Collentine, A., Meijerink, M., Menke, J., van Aert, R., Fox, J-P., Hoijtink, H., Rosseel, Y., Wagenmakers, E. J., & van Lissa, C. (2021). BFpack: Flexible Bayes Factor Testing of Scientific Theories in R. *Journal of Statistical Software, 100(18),* 1-63. DOI:10.18637/jss.v100.i18

Nosek, B. A., & Bar-Anan, Y. (2012). Scientific utopia: I. Opening scientific communication. *Psychological Inquiry*, *23*, 217–243.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ..., & Yarkoni, T. (2015). Promoting an openresearch culture. *Science, 348,* 1422–1425.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*, 615–631.

Nuzzo, R. (2014). Statistical errors. *Nature*, *506*, 150–152.

Orlitzky, M. (2012). How can signifi cance tests be einstitutionalized? Organizational Research Methods, 5, 199-228.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: Acrisis of confidence? *Perspectives on Psychological Science*, *7*,528–530.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225–237.

Sellke, T., Bayarri, M. J., & Berger, J. O. (2001). Calibration of *p* values for testing precise null hypothesis. *The American Statistician, 55*, 62-71.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني

Sharpe, D. (2013).Why the resistance to statistical innovations? Bridging the communication gap. *Psychological Methods*, *18*, 572–582.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False–positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366.

Simonsohn, U. (2015a). Posterior–hacking: Selective reporting invalidates Bayesian results also. Unpublished manuscript.

Simonsohn, U. (2015b). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, *26*, 559–569.

Tan, T. K. (2023). Informative Hypothesis for Group Means Comparison. *Practical Assessment, Research & Evaluation*, *28*, 1.

Trafimow, D., & Marks, M. (2015). Editorial. Basic And Applied

Social Psychology, 37, 1–2.

Van de Schoot, R. (2010). *Informative hypotheses. How to move beyond classical null hypothesis testing*. Utrecht, The Netherlands: Utrecht University, Ph.D. thesis. (Accessible at: http://igitur-archive.library.uu.nl/dissertations/2010-0909-200248/UUindex.html).

Vanbrabant, L. (2020). Restriktor: Constrained statistical inference (R package version 0.2-800) [Computer software manual]. https://CRAN.R-project.org/package=restriktor

Vanbrabant, L., Van de Schoot, R., & Rosseel, Y. (2015). Constrained statistical inference: Sample-size tables for

ANOVA and regression. *Frontiers in Psychology*, *5*, Article 1565. DOI: 10.3389/fpsyg.2014.01565

Vanbrabant, L., Van Loey, N., & Kuiper, R. N. (2020). Evaluating a theory-based hypothesis against its complement using an AIC-type information criterion with an application to facial burn injury. *Psychological Methods*, *25(2)*, 129-142. http://dx.doi.org/10.1037/met0000238.

Wagenmakers, E. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*, 779-804. doi: 10.3758/ BF03194105

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*, 779–804.

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic bulletin & review*, *25*, 35-57.

Wei, Z., Yang, A., Rocha, L., Miranda, M. F., & Nathoo, F. (2022). A review of Bayesian hypothesis testing and its practical implementations. *Entropy*, *24*, 161. https://doi.org/10.3390/e24020161.

Wei, Z., Yang, A., Rocha, L., Miranda, M. F., & Nathoo, F. (2022). A review of Bayesian hypothesis testing and its practical implementations. *Entropy*, *24*, 161. https://doi.org/10.3390/e24020161.