السلام عليكم و رحمة الله و بركاته

بالنسبة للمحاضرتين الأخيرة إلي شيبت فينا ، هي حبتين تحوس بس لو ركزنا على المهم ان شاء الله تسهل علينا ، فيه أشياء بيسكس و هي إلي مطالبين فيها راح أذكرها و أشرحها هنا و إنتِ عليك تحفظيها و بإذن الله يكون شيء سهل عليك حفظه ، تذكروا الدكتور عاد كم مرة مافي معادلات معانا المطلوب منك تفهمين فكرة التيست و كيف تحللين نتيجته

بالبداية مهم نفهم النول هايبوثيسس و البي فاليو ، هذا شرح سخيف يسهل عليك فهمهم للي مو فاهمتهم

(إلي فاهمة سكيب لبداية المحاضرة)

الموضوع بسيط

لما نقول نول هايبوثيسس يعني فرضية النفي يعني أنا انفي ان فيه علاقة بين شيئين تمام؟ خلاص النول هايبوثيسس هي فرضية النفسيات الي يشوفوا كل شيء مستحيل و غير متحقق

طيب أنا لو أبغى أبحث ان أكل البطاطس المقلية مع الكتشب يزيد رغبة الانسان في أكل المزيد من البطاطس

النول هايبوثيسس حقتي لهالبحث تنفي هالشيء

تقولك والله حطيتي كتشب ماحطيتي بتاكلين نفس الكمية

طيب بالبحث لو جبنا مجموعة ناس و قلنا لهم ها يا عيال الكتشب يحمسكم تاكلون بطاطس زيادة او لا؟ و طلع جواب اغلبهم ايوة

هل اقدر اقول هههه نظريتي صح و فعلاً الكتشب يزيد من اكلنا للبطاطس

لا حبيبتي الموضوع مو بهالسهولة

عادي بيجي واحد يقول تستهبلين ؟؟؟؟ هذا بحثك على عينة و امكن صدفة سألتي ناس يحبون الكتشب و الحقيقة ان مافي علاقة

هنا انت ايش تسوين يا ربيييييييي؟؟؟؟؟ البي فاليو ياحلوة

هي الي بتثبت لك هل فعلاً كان نتيجة بحثك صدفة او لا

شلون هالبي فاليو ؟؟؟

"البي فاليو هي تحسب لك "احتمالية ان هالعلاقة الي طلعت لك صدفة

لو طلعت هالاحتمالية قليلة جداً

معناته والله بحثك بطل و تقدرين بكل فخر تردين عليه و تقولين له احتمالية انه صدفة (البي فاليو) قليلة جداً

و بالتالي انا ارفض كلامك الي هو مافي علاقة بين الكتشب و اكل البطاطس (النول هايبوثيسس)

و بناءً على ذلك أميل لتصديق ان فيه علاقة

فكذا تكوني حجرتي له بطريقة علمية

وهو كذا الوضع مع الأبحاث

ما اقدر اختبر صحة فرضيتي بشكل مباشر

يعني ما أقدر اختبر الألتيرنات هايبوثيسس

بس أقدر أشوف احتمالية الصدفة

اذا قليلة معناته ارفض فرضية النفي النول هايبوثيسس

يعني ارفض انه مافي علاقة

و بكذا ارجح ان فيه علاقة

يعني شوي جحا لان ماتقدرين تثبتين بشكل مباشر

بس هو كذا كأن الاصل في الاشياء ان مابينهم علاقة

فانتي عشاني تثبتين ان فيه علاقة لازم

تطبقين الدراسة بشكل صحيح و عشوائي (١

تحصلين على نتائج ايجابية تتفق مع كلامك (٢

تثبتين ان نتائجك هذي مو صدفة عن طريق حساب احتمالية الصدفة (٣

لازم تطلع احتمالية صغيرة جداً (٤

وقتها ترفضين ان مافي علاقة (٥

طيب المحاضرة بسم الله بعنوان

Statistical tests for Quantitative variables (z-test, t-test & Correlation)

هي محاضرة خفيفة (يا رب تصدقوني)؛ ماعندك إلا ثلاث اختبارات الزي و التي و الكوريلايشن

أول شيء مهم تعرفين ايش هذي الاختبارات؟ اختبارات احصائية تأكد إن نتيجتك ليست محض صدفة و تنفي النول هايبوثيسس (لو فزتي و كان نتيجة الاختبار ايجابية) ، طيب حسبي الله عليهم ليش فيه اختبارات كثير خلاص يحطون اختبار واحد و يريحونا : شوفي احنا عندنا دراسات كثيرة و متنوعة بطريقة عمل الدراسة و هدفها و فكرتها و كل شيء فلازم يعني يكون عندنا اختبارات مختلفة و لذلك راح نتعرف على كل اختبار و متى نستخدمه و كيف نفسر نتيجته بس عموماً الفكرة متشابهة مرة فلا تحسين إن الموضوع ثقيل

بسم الله نبدأ أول شيء

Z-test: normal distribution test

Study variable: Qualitative (Categorical)

Outcome variable: Quantitative, Qualitative

Comparison:

(i)sample mean with population mean

(ii)two sample means

Sample size: larger in each group (>30) & standard deviation is known

حطي ببالك هذا التيست هو تبع النورمل ديستربيوشن فمعناته يحتاج سامبل سايز كبير ، لأن زي ما عرفنا من محاضرة النورمل ديستربيوشن ان الدراسة كل ما زاد حجم العينة فيها كانت أقرب للنورمل ديستربيوشن فبالتالي هنا مهم تعرفين شرط الزي تيست عينة أكبر من 30 ، لو عينة أقل من ثلاثين غير ممكن تطبقين عليها الاختبار هذا ، لاحظي الاوت كوم فاريبل الاثنين كوانتييتف و كوالتيتتف

هذا الاختبار يستخدم في حال كانت الدراسة للمقارنة بين متوسط السامبل و متوسط البوبيولايشن مثل : لو اخذنا مثل متوسط دخل الفرد السعودي على مستوى المملكة (البوبيوليشن كله) من قسم الاحصاء التابع للدولة و كان عشر الالف جا واحد و قال والله هممم ما احس الكلام هذا صح و قرر ياخذ عينة من كل مدينة و حسب متوسط دخلهم و طلع 15 الف

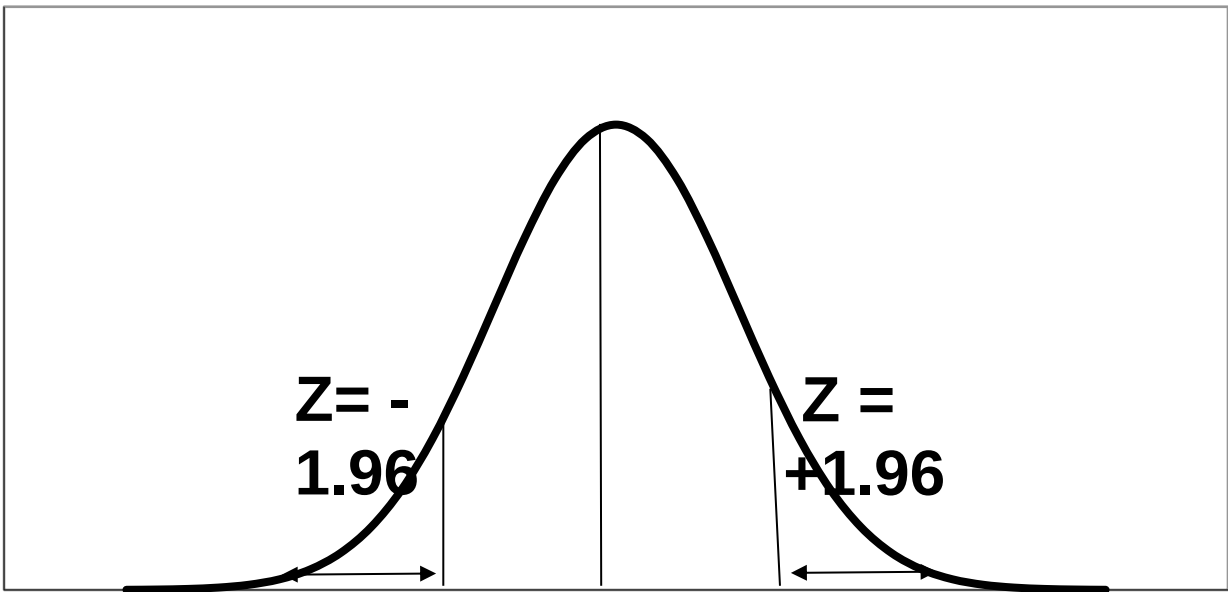
فنستخدم هذا الاختبار عشان نتحقق هل هو حصل على ال15 الف صدفة و لا فعلا فيه فرق زي مايقول

و أيضاً يستخدم للمقارنة بين عينة و عينة يعني مثلاً لو رجعنا لمثال الدخل مايكون كذا يقارنه مع البوبيوليشن لكن مثلاً يقول بقارن دخل الي بالدمام و دخل الي بالرياض و ياخذ عينة من الدمام و عينة من الرياض و يقارن بينهم و بعد ماتطلع النتيجة يتحقق عن طريق هذا الاختبار تمام؟

طيب كيف طريقة الحل و فهم الاختبار هذا؟ و لله الحمد المعادلات مو معانا فماراح أذكرها لكن عموماً بقولكم الطريقة احفظوها

الآن أول شيء نحسب الميين (المتوسط) لكل عينة أو للعينة و للبوبيولايشن (هم بيذكروا الارقام بالسؤال) غالباً حيكون بينهم فرق يعني الأرقام مختلفة لكل عينة طيب هنا يجي دور الزي تيست هل هذا الفرق حقيقي أو فقط بسبب الصدفة طلع لنا؟؟؟

الزي تيست نحسب فيه الألفا 0.05 دائماً لأنها أقل قيمة مقبولة ، لو طبقنا قيمة الألفا على النورمل ديستربيوشن لما كنا نقسمه بيطلع لنا الزي فاليو +1.96 من اليمين و -1.96 من اليسار لأنه (تو سايدد) ، مو فاهمين هذا الكلام؟؟؟ هو مبني على المحاضرات السابقة و لو شفتوا الصور بتفهمون شوفوا تحت الصورة و الكلام



عموماً كل إلي تحتاجوا تعرفوه إنها 1.96 موجب من اليمين و 1.96 سالب من اليسار هذي حدود ايش؟ هذي حدود منطقة قبول النول هايبوثيسس ، يعني لو كانت القيمة (النتيجة الي تطلع لنا لما نطبق معادلة اختبار الزي على بحثنا) بين هذي الفترة و في هذي المنطقة يصير نقبل النول هايبوثيسس و نقول والله مافي فرق و لا نقدر نقول الفرق الي طلع كان صدفة و البحث كلام فاضي ، لكن لو حسبنا و طلع و الله القيمة أكبر من +1.96 و لا أصغر من -1.96 نقول وقتها لا و الله هذي منطقة (رفض) النول هايبوثيسس يعني و الله فيه الفرق ما كان فرق صدفة كان فرق حقيقي

خلاص ؟ هذا أهم شيء تعرفيه بالتيست ، ارجعي للأمثلة تبع الدكتور و شوفي ايش كانت النتايج و فسريها إنتِ أول بناء على كلامنا و لو طلع صح ، برافو كملي الحين

فيه ملاحظة على البي فاليو اذا كان الكلام على سامبل و بوبيوليشن مانقدر نحدد قيمة بس نقدر نقول ان شرطنا البي فاليو اقل من 0.05 ، لكن لو كانت عينة و عينة يعني البل مقفول من الجهتين نقدر نحسب مناطق رفض النول هايبوثيسس (الي بالأطراف) وهي من محاضرة البي فاليو و بتكون تساوي 0.03 (من الجهتين 0.015)

نجي للتي تيست

Student's t-test:

Study variable: Qualitative (Categorical)

Outcome variable: Quantitative (mean values only)

Comparison:

(i)sample mean with population mean

(ii)two means (independent samples)

(iii)paired samples

Sample size: each group <30 (can be used even for large sample size)

نلاحظ هنا الآوت كوم فاريبل بس كوانتيتيتف ، نسخدمه في ثلاث حالات سامبل و بوبيوليشن مثل الزي تيست و تو سامبل مختلفين عن بعض نقول عليهم ان ديبندنت ، و الشيء الزيادة هنا الي يميز التي تيست البيرد سامبلز الي هي ديبيندنت سامبلز ، هنا ما راح ناخذ عينتين مختلفة و نقارنهم لا لا ، نفس العينة بس قبل و بعد التدخل (الانترفينشن) يعني مثلاً أجيب مجموعة ناس و أعطيهم اختبار بشيء ما قد درسوه بعدين بعد ما ادرسهم و اعلمهم اختبرهم (نفسهم) مرة ثانية و أقارن نتيجتهم قبل ما يدرسون و بعد ما درسوا هذا الديبيندنت او البيرد سامبل ، كمان نلاحظ ان التي تيست كتكوت و متواضع مايتشرط عادي عنده سامبل سايز قليل عادي اقل من 30 مو زي الزي تيست

طيب هنا عندنا جدول للتي مثل الزي ، التي لما نجي نسميها نطرح من حجم العينة 1 يعني مثلاً عندي سامبل سايز 9 معناته التي حقهم 8 فأروح افتح جدول التي و اشوف قيمة التي 8 عند ليفل سيقنيفكنس (الي تعودنا عليه 0.05) طبعاً بالجدول عندنا ون سايدد و تو سايدد بقولكم الفرق بعد شوي ، عموماً يعني زبدة الموضوع مافي جدول بالاختبار بس يعني عشان تفهمين السالفة راح نطلع قيمة التي من الجدول نسميها تيبل تي 8 ، و بالمعادلة الخاصة بهذا الاختبار نحسب تي 8 الي هي نسميها كالكولايتد تي 8 بعد كذا نقارن : لو الكالكولايتد تي أكبر من التيبل تي معناته أرفض النول هايبوثيسس و الله البحث قوي و فعلاً الفرق حقيقي و مو صدفة ، لكن لو الكالكولاتيد تي أصغر من التيبل تي هه ماش لازم أقبل النول هايبوثيسس و أقول الفرق كان صدفة و الأصل مافيه فرق يا بابا

^ هذا كيف نفسر نتيجة التي تيست و هذا أهم شي تعرفيه

طيب متى نقول تو سايدد و ون ساید ؟ تو سايدد في معظم الحالات لأننا في الغالب مانعرف هل هذا اقل و لا اكثر و لا يزيد و لا ينقص فممكن يكون من الناحيتين السالبة او الموجبة فنستخدم تو سايدد ، لكن لكن لكن

When the researcher knows that the intervention is going to give positive result it's one-sided but if you're not sure it will be 2 sided

زي حالة البيرد سامبل (مو دايماً) بس يعني انا عارفة لو تعلموا و درسوا راح يحلون بالاختبار بشكل أحسن فمتوقعة درجاتهم تزيد يعني من الناحية الايجابية فهنا يكون ون سايدد

كذا خلصنا الزي و التيست فنقول باختصار:

"Z and t" are the measures of:

How difficult is it to believe the null hypothesis?

High z & t values

Difficult to believe the null hypothesis -
accept that there is a real difference.

Low z & t values

Easy to believe the null hypothesis -
have not proved any difference.

يعني لما نطبق المعادلات تبعهم وتطلع النتيجة (احنا تجينا النتيجة مكتوبة) كبيرة معناته والله نرفض النول هايبوثيسس و نقول فيه فرق حقيقي و لما النتيجة قليلة ماراح نقدر نرفض النول هايبوثيسس يعني راح نقبلها و ماراح نقدر نثبت ان فيه فرق

نجي لآخر شيء الشيء الخفيف اللطيف الكوريلايشن ، ايش الكوريلايشن؟

A number called the correlation measures both the direction and strength of the linear relationship between two related sets of quantitative variables.

هم شيئين بينهم علاقة خطية بينهم رابط ، مثلاً اجتهادك و مذاكرتك و درجتك بالاختبار بينهم علاقة ؟ ايوة ، ايش هي العلاقة اذا زاد واحد زاد الثاني؟ ايوة معناته علاقة خطية هنا نقدر نطبق الكوريلايشن حلوين ؟؟؟؟

عندنا نوعين من الكوريلايشن ايجابية و سلبية

Positive – Variables move in the same direction

Examples: Height and Weight, Age and BP

الايجابية الاثنين زي بعض لو زاد هذا زاد الثاني ، لو نقص هذا نقص الثاني و هكذا

Negative Correlation - Variables move in opposite direction

Examples: Duration of HIV/AIDS and CD4 CD8, Price and Demand, Sales and advertisement expenditure

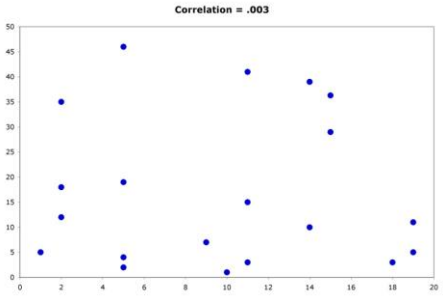
هنا السلبية العكس لو زاد واحد الثاني ينقص مثل الاتش آي في لو زادت حدة الاصابة فيه تقل الخلايا المناعية سي دي 4 ، أو مثلاً الأسعار كل ما كان الشيء غالي كل ما قل عدد الناس الي يشتروه و هكذا

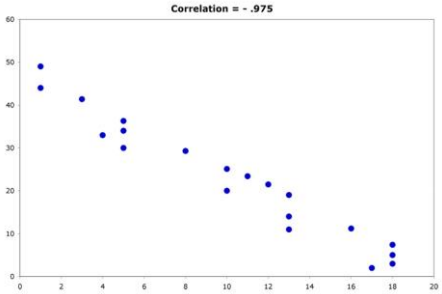
طيب هذي العلاقة تختلف حدتها من شيء لشيء كيف نختبرها ؟

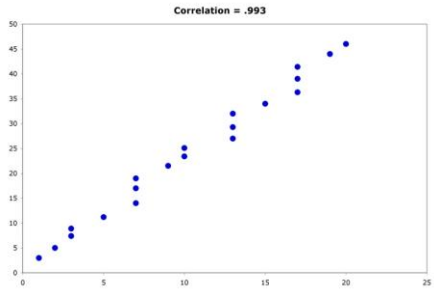
- **Measurement of correlation**

1. **Scatter Diagram**
2. **Karl Pearson's coefficient of Correlation**

شوفوا السكاتير داياقرامز







No correlation **Negative Correlation** **Positive correlation**

و عندنا كمان البيرلسون فيه معادلة ماعلينا فيها بس يعني نحسب منها الكو افيشنت الي هو آر كل ما كانت القيمة قريبة من 1 معناته هذي علاقة إيجابية قوية و كل ما كانت قريبة من -1 معناته علاقة سلبية قوية و كل ماكانت قريبة من الصفر معناته ماكو علاقة

الحين اقروا الكلام هذا لأن مهم

a correlation coefficient (r) provides a quantitative way to express the degree of <u>linear</u> relationship between two variables.

- **Range: r is always between -1 and 1**
- **Sign of correlation indicates <u>direction</u>:**

- high with high and
low with low -> positive

- high with low and
low with high -> negative

- no consistent pattern -> near zero

- **Magnitude (absolute value) indicates <u>strength</u> (-.9 is just as strong as .9)**

.10 to .40 weak

.40 to .80 moderate

.80 to .99 high

1.00 perfect

فيه أشياء نحتاج نعرفها عن الآر

- **r is not dependent on the units in the problem (different variables) عندنا المتغيرين غير عن بعض في العادة فعادي يعني لو اختلفت الوحدات مثلاً العمر و الوزن**
- **r ignores the distinction between explanatory and response variables doesn't matter**

لما نتكلم عن علاقة بين شيئين عندنا متغير اساسي و متغير تابع مثلاً العمر و الوزن نعرف ان العمر كل مازاد زاد الوزن هالاختلاف بين المتغيرين الاكسبلاناتوري و الريسبونس فيهم مايهمنا في الآر

- **r is not designed to measure the strength of relationships that are not approximately straight line ما نقدر نقيس الا العلاقات الخطية البحتة غير كذا لا**
- **r can be strongly influenced by outliers (extreme value)**

هنا بعد نقاط مهمة أقروها

1. **Correlation coefficient is appropriate measure of relation <u>only when relationship is linear</u>**
2. **Correlation coefficient is appropriate measure of relation when equal ranges of scores in the sample and in the population.**
3. **Correlation <u>doesn't imply causality</u>**

- **Using U.S. cities a cases, there is a strong positive correlation between the number of churches and the incidence of violent crime**
- **Does this mean churches cause violent crime, or violent crime causes more churches to be built?**
- **More likely, both related to population of city (3d variable -- lurking or confounding variable)**

هنا يقولك حتى لو كان عندك علاقة خطية ماتقدرين تقولين هذا سبب الثاني يعني مثلاً لو قلنا كل ما زاد بيع الايس كريم زادت الجرايم معناته الايس كريم سبب الجرايم ؟؟؟؟ لا بس لان زيادة البيع تدل على متغير ثالث كونفاوندينق الي هو زيادة عدد الناس و بالتالي تزيد الجرايم فهنا مانقدر نقول تفسيرات سببية اذا ما حللنا الموضوع بشكل صحيح

In conclusion !

Z-test will be used for both categorical(qualitative) and quantitative outcome variables.

Student's t-test will be used for only quantitative outcome variables.

Correlation will be used to quantify the linear relationship between two quantitative variables

يا رب فهمتوا المحاضرة و طلعت فعلاً أخف مما تتوقعون

المحاضرة الي بعدها تشبهها من حيث الفكرة لكن الفرق ان هناك نبغى نختبر كوالتيتيتف ستوديز فراح نستخدم اختبارات مختلفة أول اختبار الزي تيست لان ذكرنا يستخدم للاثنين كوالتيتيتف و كوانتيتيتف الفرق اننا مانستخدم المبين نستخدم البروبوشن و الباقي تقريباً نفس الشيء ، لاحظوا بس عندنا الكي سيكوير مانستخدم بروبوشن بالجدول نفسه لما نسويه نستخدم فريكونسيز عدد مرات ظهور المتغير مرة كان ودي أشرح المحاضرة إلي بعدها بس للأسف مافيه وقت ، إذا فهمتوا هذي المحاضرة و فهمتوا ايش المطلوب منكم بإذن الله تسهل عليكم المحاضرة الي بعدها ، ركزوا على تفسير كل اختبار متى نقبل النول و متى نرفضها ، إلهي يفتح عليكم و يلهمكم الصواب

دعواتكم

جوهرة المالكي