علم الإحصاء

لا يقدم علم الرياضيات نماذج للظواهر المتوقعة فقط، كما هو الحال في علم الميكانيكا النيوتونية Newtonian Mechanics، بل يقدم أيضاً أدوات ثمينة لا تقدر بثمن لتحليل مبدأ الشك uncertainity. وهذا المجال يسمي الاحتمال Probability. وفي هذا المجال بإمكان التجارب البسيطة مثل إلقاء حجر النرد، أو إلقاء قطعة النقود أن تقدم رؤى نافذة في مجال رياضة الاحتمالات.

لا يستغنى أي عمل أو أي حكومة في العالم عن استخدام علم الرياضيات متمثلا في الإحصاء استخداما يوميا؛ فهو الأداة المستخدمة لتحليل كل أشكال البيانات العددية.

لا عجب في أن علمي الاحتمالات والإحصاء يرتبطان ببعضهما البعض ارتباطا وثيقا، فالبيانات الإحصائية التي تجمع من خلال التجارب تقدم دلائل حول توزيع الاحتمال الكامن وراءها، وكذلك يمكن إثبات عديد من النتائج الضخمة المنسجمة لمثل هذه التوزيعات مثل: قانون الأعداد الكبيرة the law of large numbers ، ونظرية الحد المركزية the central limit

theorem، فمن الصادم أن هذه الحسابات الرياضية غالبا ما تتعارض بشدة مع الحدس البشري. ومن أمثلة ذلك: مسألة مونتي هول Monty Hall problem، ومغالطة المدعي the prosecutor`s fallacy. (ويرجح البعض أن هذه السمة قد تكون متأصلة في التطور). ويقوم الناس بتطبيق الاستدلال البايزي techniques of Bayesian inference لتعزيز قدراتهم على تقدير حجم المخاطرات ومحاربة الميل إلى اللاعقلانية.

يعد علم التشفير Cryptography فرع مهم من فروع الرياضيات التطبيقية. ويتراوح هذا العلم مابين القدم والحداثة، فهو يضم أقدم وسيلة لتشفير الرسائل وهي التشفير أحادي الحرف monoalphabetic cipher، كما يضم نظرية المعلومات؛ تلك النظرية الحديثة التي تدعم الإنترنت.

الإحصاء STATISTICS

المتوسط Mean

لنفرض أننا خرجنا وقمنا بجمع مجموعة البيانات الآتية:

3, 3, 4, 3, 4, 5, 3, 9, 4

سنطلق على هذه المجموعة A. قد تكون هذه المجموعة هي أعداد أوراق، أو أعداد السكان في شارع ما. مهما كان الشيء الذي تمثله هذه الأرقام فإننا قد نحتاج حساب معدلها. المتوسط الحسابي the mean، والوسيطmedian، والمنوال mode ، ومتوسط المدى mid-range جميعها تعبر عن المعدل average ولكن بصيغ حسابية مختلفة وغالبا لا تكون نواتج هذه الصيغ متساوية.

المتوسط mean هو أشهر صيغ حساب المعدل Average ويحسب بجمع الأرقام التي نريد حساب متوسطها (3 + 3 + 4 + 3 + 4 + 5 + 3 + 9 + 4 = 38)، ثم القسمة على عددها (9 في هذا المثال)، وبالتالي يكون المتوسط في هذا المثال هو $\frac{38}{9} = 4.2$ 4.2 =38/9.

وبشكل عام إذا كان لدينا عدد m من الأرقام:a1, a2, ... , am فيمكن حساب متوسطها من العلاقة: $\frac{a_1+a_2+\cdots+a_m}{m}$. ويمكن أن يأخذ المتوسط تسلسلات غير متوقعة، فعلى سبيل المثال: متوسط عدد أذرع البشر حول العالم لا يساوي 2؛ إنه يساوي تقريبا 1.999 (معظم البشر لهم ذراعان، وبعضهم إما أن يكون بلا أذرع أو له ذراع واحد، وعدد قليل جدا لديه ثلاثة أذرع أو أكثر). بالتالي الأغلبية الساحقة من البشر لديهم عدد أذرع يفوق المعدل وفي هذه الحالة نستخدم المتوسط mean.

الوسيط Median

لنفرض أن لدينا مجموعة A من البيانات.3, 3, 4, 3, 4, 5, 3, 9, 4 ونريد حساب الوسيط–median فإننا نقوم أولا بترتيب القيم ترتيبا تصاعديا :

3, 3, 3, 3, 4, 4, 4, 5, 9

فيكون الوسيط هو الرقم الذي يقسم الأرقام إلى نصفين وهو العدد 4 في هذا المثال.

ولكن يأتي التعقيد عندما يكون عدد البيانات رقما زوجيا، ففي هذه الحالة يكون الوسيط هو المتوسط الحسابي للرقمين الذين يقعان في المنتصف. على سبيل المثال: لنفرض أننا نريد حساب الوسيط لمجموعة القيم : 9,10,12,14 . لايوجد بين تلك القيم رقما يقع في المنتصف (لأن عدد القيم زوجي)، لكن يوجد قيمتين في المنتصف هما : 10,12؛ لذلك نحسب المتوسط لهما فيكون الناتج هو الوسيط $\frac{10+12}{2} = 11$.

المنوال Mode و متوسط المدى mid-range

المنوال هو القيمة الأكثر تكررا في مجموعة بيانات. في المثال السابق، منوال المجموعة A هو القيمة 3. بعض مجموعات البيانات يكون المنوال ليس له معنى مثل: 1,2,3,4,5 أو 2,2,50,1001,1001.

أما المدى فهو أبسط صور المعدل؛ فهو المتوسط الحسابي لأكبر وأصغر قيمة. في المثال السابق، متوسط المدى للمجموعة A هو $\frac{3+9}{2} = 6$.

الجداول التكرارية Frequency tables

كان التعامل مع المجموعة A في الأمثلة السابقة سهلا؛ فهي مكونة فقط من 9 قيم، ولكن في الكثير من التطبيقات العملية تكون مجموعات البيانات أكبر بكثير.

وتظل الأساليب الرياضية المتبعة في حساب المتوسط الحسابي، والوسيط الحسابي، والمنوال كما هي سواء أكانت مجموعات البيانات كبيرة أم صغيرة، يكون هناك فقط اختلاف طفيف في تمثيل البيانات.

لنفرض أن لدينا بيانات عن عدد سكان كل منزل أو شقة في مدينة ما فإنه يمكننا تمثيل النتائج فيما يسمي الجداول التكرارية:

عدد السكان في العقار	عدد مرات التكرار
0	292
1	5745
2	8291
3	4703
4	2108
5	961
6	531
22,631	المجموع

هذا يعنى أنه يوجد 292 عقار خال من السكان، و5745 عقار يسكنه فرد واحد، وهكذا. ومجموع العقارات في المدينة يساوي 22,631 عقار. وأسهل معدل يمكن حسابه هنا هو المنوال الذي يمكن استخراجه مباشرة من الجدول: وهو هنا = 2 لأنه هو القيمة الأكثر تكرارا من بين البيانات.

حساب المتوسط باستخدام الجداول التكرارية Mean from frequency tables

لحساب المتوسط mean من الجدول التكراري السابق نقوم بحساب مجموع عدد سكان جميع العقارات ثم نقسم هذا المجموع على عدد العقارات. يحتوى عمود "عدد مرات التكرار" على عدد المنازل ومجموعها يساوي 22,631 منزلا، ولكن كيف يمكن حساب مجموع عدد السكان؟

من الواضح أن عدد سكان المنازل الخالية من السكان يساوي صفرا. ويوجد 5745 عقار يشغله ساكن واحد وهذا يسهم بعدد 5745 ساكن من مجموع السكان. وعدد السكان الكلي في العقارات التي يسكنها فردان هو 8291×2 =16,582 ساكن. وعدد السكان الكلي في العقارات التي يسكنها ثلاثة أفراد هو 4703×3 = 14109ساكن.

بذلك يتضح أنه لحساب العدد الكلي للسكان نقوم بحساب حاصل ضرب أول عمودين في الجدول. للتسهيل سوف نعطي العمود الأول الرمز n، ونرمز لعدد مرات التكرار بالرمز f ، ثم نضيف عمودا جديدا للجدول لوضع حاصل ضرب n x f.

n	f	n x f
0	292	0
1	5745	5745
2	8291	16,582
3	4703	14,109
4	2108	8432
5	961	4805
6	531	3186
المجموع	22,631	52,859

يمثل العمود الثالث (n x f) العدد الكلي للسكان في كل منزل. وبجمع هذا العمود نحصل على العدد الكلي للسكان في المدينة ويساوي 52,859 وبذلك يمكننا أخيرا حساب المتوسط الحسابي: $\frac{52.859}{22.631} = 2.34$ مقربا لأقرب رقميين عشريين).

ويمكن كتابة المتوسط الحسابي بطريقة أخرى باستخدام رمز المجموع

$$\frac{\sum n \times f}{\sum f}$$

التكرار التراكمي Cumulative frequency

الوسيط median في الجدول التكراري السابق هو العقار الذي يقع في منتصف القائمة وهو العقار الذي ترتيبه 11316. لذلك نريد أن نستنتج الفئة المناسبة. الطريقة المناسبة لعمل ذلك هي إضافة عمود لحساب التكرار التراكمي.

عدد السكان في العقار	عدد مرات التكرار	التكرار التراكمي
0	292	292
1	5745	6037
2	8291	14,328
3	4703	19,031
4	2108	21,139
5	961	22,100
6	531	22,631

المدخل الأول في عمود التكرار التراكمي هو 292، وهو نفسه المدخل الأول في عمود التكرار البسيط. أما المدخل التالي فهو مختلف؛ حيث إن عدد السكان في العقار الذي يشغله ساكن واحد يساوي 5745 ساكن بينما 6037 هو عدد السكان في العقارات الخالية من السكان أو التي يشغلها ساكن واحد ، وبالمثل المدخل الثالث 14328 هو عدد العقارات الخالية من السكان أو التي يشغلها ساكن واحد أو ساكنان.

التكرار التراكمي في أغلب الأحيان يكون مفيدا، خاصة أنه يسهل عملية حساب الوسيط الحسابي . ففي هذا المثال نريد أن نبحث عن المنزل الذي ترتيبه 11316، وبما أن فئة المنزل الذي يشغله ساكنان يشمل المنزل الذي ترتيبه 6038 حتى المنزل الذي ترتيبه 14328 إذا لابد أنه ينتمي إلى هذه الفئة،وبالتالي الوسيط الحسابي يساوي 2.

المدى الربيعي Interquartile Range

المعدلات المختلفة لمجموعة من البيانات هي وسائل لتحديد مركز هذه المجموعة. ومن الوسائل الأخرى المهمة في هذا الصدد تشتت القيم spread وتوجد طرق مختلفة لوصفه.

المدى range هو الفرق بين أكبر قيمة وأصغر قيمة في البيانات المجمعة، وهذا لا يقدم قياسا مفيدا؛ لأن اهتمامه ينصب في الأساس على القيم المتطرفة (النقاط التي تقع بعيدا عن الإطار العام الذي ندور في فلكه)، وهنا يظهر مقياس أخر هو الانحراف الربيعي.

يحسب الوسيط عن طريق ترتيب كل القيم ترتيبا تصاعديا ثم تكون القيمة التي تقع في منتصف المسافة تماما بين القيمة العظمى والصغرى هي الوسيط الحسابي. وبالطبع يمكن النظر إلى النقاط التي تقع في ربع أو ثلاثة أربع المسافة بين هاتين النقطتين، ويطلق على تلك النقاط اسم الربيعات–quartiles ، وتسمى المسافة بينها الانحراف الربيعي– Interquartile Range. فعلى سبيل المثال: إذا كان لدينا مجموعة من البيانات لها القيم الآتية : 1, 1, 3, 5, 7, 9, 10, 15, 18, 20, 50 فيكون الوسيط الحسابي هو 9، والربيعات

هي 3، و 18 فيكون الانحراف الربيعي مساويا لـ 18-3=15 والمدى الكلي 50-1= 49 . أما إذا كان لدينا مجموعات بيانات كبيرة فإننا نستخرج الربيعات من جدول التكرار التراكمي:

عدد السكان في العقار	عدد مرات التكرار	التكرار التراكمي
0	292	292
1	5745	6037
2	8291	14,328
3	4703	19,031
4	2108	21,139
5	961	22,100
6	531	22,631

الربيع الأول هو المنزل الذي ترتيبه 5658 الذي يقع في فئة المنازل التي يشغلها ساكن واحد.الربيع الثالث هو المنزل الذي ترتيبه 16947 الذي يقع في فئة المنازل التي يسكنها ثلاثة أفراد ويكون الانحراف الربيعي هو 3-1=2، والمدى الكلي=6-0=6.

تباين العينة sample variance

يتشابه تباين العينة مع الانحراف الربيعي في أنه هو أيضاً أحد مقاييس التشتت. وتباين العينة هو المقياس الطبيعي الذي ننظر إليه بعين الاعتبار عندما نتعامل مع الوسيط على أنه مركز العينة. لنفرض مجموعة X من البيانات تحتوي على عدد n من النقاط، وأن الوسط هو μ، يمكن حساب المسافة بين أي نقطة x ونقطة الوسيط من العلاقة $x-\mu$ ، ثم نقوم بتربيع الناتج لنحصل على قيمة موجبة $(x-\mu)^2$.

وإذا قمنا بذلك مع جميع قيم X ثم أخذنا الوسيط للنواتج نكون قد حسبنا التباين- variance للمجموعة X ونطلق عليه VarX ويحسب من العلاقة

$$\text{Var X} = \frac{\sum(x-\mu)^2}{n}$$

على سبيل المثال:

لدينا مجموعة من البيانات لها القيم الآتية: 3,3,3,3,4,4,4,5,9 يكون متوسطها 4.2، ولحساب التباين نطرح قيمة المتوسط من كل قيمة على حدة فتكون القيم بعد الطرح كالتالي:

1.2- ، 1.2- ، 1.2- ، 1.2-،0.2-،0.2-،0.7 ،4.7

ثم نقوم بتربيع القيم بعد الطرح وحساب متوسطها

$$\frac{4 \times (-1.2)^2 + 3 \times (-0.2)^2 + (0.7)^2 + (4.7)^2}{9} = 3.284$$

(لأقرب ثلاثة أرقام عشرية)

ويكون الانحراف المعياري standard deviation للمجموعة X مساويا $\sqrt{\text{VarX}}$ وهو يساوي في المثال السابق $\sqrt{3.284} = 1.812$ (لأقرب ثلاثة أرقام عشرية). (انظر التباين و القيم المتوقعة للتوزيع الاحتمالي).

العزوم Moments

الطريقة التي ذكرناها لحساب التباين صحيحة وتعمل جيدا ولكن يوجد طريقة أسرع قليلا، فقد وجد أن التباين يساوي المتوسط الحسابي لمربعات القيم مطروحا منها مربع المتوسط

$$\text{Var X} = \frac{\sum x^2}{n} - \mu^2$$

نقوم بحساب متوسط مربعات مجموعة البيانات التي قيمها: 3,3,3,3,4,4,4,5,9 :

21.1=9/4*23+3*24+25+29 دائر

ثم نطرح مربع المتوسط

21.1–$(4.2)^2$ =3.284 (لأقرب ثلاثة أرقام عشرية)

يعرف المتوسط $\frac{\Sigma x}{n}$ أيضاً باسم العزم الأول the first moment، أما متوسط مربعات

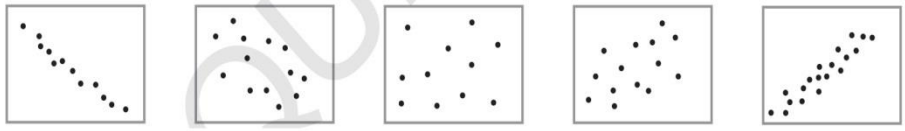
القيم $\frac{\sum x^2}{n}$ فيعرف باسم العزم الثاني second moment، ويستخدم علماء الإحصاء عزوما ذات رتب أعلى في تحليلاتهم.

الارتباط Correlation

هل تزيد إحتمالية إصابة المدخنين بمرض السرطان عن غيرهم؟ هل عدد أبناء سكان البلاد الممطرة أكبر من عدد أبناء سكان البلاد الأخرى؟ هل الخنافس الأكثر وزنا تتمتع بأعمار أطول؟ يحتاج العلماء في كثير من الأحيان إلى دراسة العلاقة بين ظاهرتين، والأداة الإحصائية التي تقوم بذلك هي الارتباط

لنفرض أن لدينا بيانات عن وزن بعض الخنافس وأعمارها وأردنا اختبار إذا ما كانت هناك علاقة بين الوزن والعمر أم لا. يمكننا البدء بعمل رسم بياني بين الوزن والعمر، فإذا كانت النقاط الناتجة تبدو عشوائية ومشتتة إذا لا يوجد ارتباط بين عاملي الوزن والعمر، وعلى النقيض من ذلك، إذا كانت ترتيب النقاط أشبه ما يكون بالخط المستقيم إذا العاملان يرتبطان ارتباطا قويا وما يقع بين الحالتين السابقتين يكون ارتباطا ضعيفا.

بفرض وجود نوع من الارتباط، إذا كان الزيادة في الوزن يصاحبها زيادة في طول العمر يسمي ارتباط طردي، أما إذا كان طول العمر يقل بزيادة الوزن يسمى الارتباط ارتباط عكسي



هناك العديد من الأساليب التي يمكن لعلماء الإحصاء من خلالها تعيين رقم يصف الارتباط ، ويسمي هذا الرقم معامل الارتباط–correlation coefficient (ومنها معامل ارتباط الرتب لسبيرمان–Spearman's rank correlation coefficient)

وفي جميع الأحوال، تتراوح قيمة معامل الارتباط بين -1 ، +1 بحيث تعبر النتائج

القريبة من -1 عن ارتباط عكسي قوي، والنتائج القريبة من +1 عن ارتباط طردي قوي، وإذا كانت النتيجة صفرا دل ذلك على عدم وجود ارتباط. وحتى إذا كشفت دراستنا لعاملي وزن وطول عمر مجموعة من الخنافس عن وجود ارتباط طردي بين هذين العاملين فليس بإمكاننا أن نعرف إذا ما كانت الخنافس الأكثر صحة، والأطول عمرا تكون زائدة الوزن أم أن الوزن الزائد يساعد على حماية الخنفساء من التعرض للإصابات أم أن هناك عامل ثالث لم ندرسه مثل النوع: فقد تكون أنثى الخنفساء أثقل وزنا وأطول عمرا من ذكر الخنفساء ، وهنا يجب توخي الحذر لأن الارتباط لا يتضمن عنصر السببية (لا يساعد على معرفة أي العوامل سبب للآخر).

معامل ارتباط الرتب لسبيرمان Spearman's rank correlation coefficient

توجد طرق عديدة لاختبار الارتباط. ويتميز معامل ارتباط سبيرمان للرتب بأنه لا يشترط أن يكون الارتباط ارتباطا خطيا حتى يمكن تطبيقه.

بفرض أن العلماء قد اكتشفوا نوعا جديدا من النباتات المضيئة ، وأرادوا البحث في إذا ما كان هناك ارتباط بين طول النبات ودرجة إضائته . سوف أقوم بأخذ عينة صغير لتوضيح هذه الطريقة: .

النبات	طول النبات	درجة الإضاءة
A	6.1	0.41
B	4.5	0.37
C	5.0	0.36
D	5.9	0.31
E	7.3	0.45
F	6.2	0.38

الخطوة الأولى هي ترتيب النباتات حسب الطول ودرجة الإضاءة (بحيث تكون الرتبة 1 هي رتبة النبات الأطول والأكثر إضاءة).

النبات	طول النبات	رتبة طول النبات	درجة الإضاءة	رتبة درجة الإضاءة
A	6.1	3	0.41	2
B	4.5	6	0.37	4
C	5.0	5	0.36	5
D	5.9	4	0.31	6
E	7.3	1	0.45	1
F	6.2	2	0.38	3

الآن يمكننا أن ننحي البيانات الحقيقية جانبا ، ونعمل على الرتبتين. ثم نحسب فرق الرتب (d) لكل نبات ونحسب مربعه (d^2).

النبات	رتبة طول النبات	رتبة درجة الإضاءة	فرق الرتب (d)	مربع فرق الرتب (d^2)
A	3	2	1	1
B	6	4	2	4
C	5	5	0	0
D	4	6	-2	4
E	1	1	0	0
F	2	3	-1	1

ثم نجمع عمود مربع فرق الرتب (d^2) الذي مجموعه في المثال السابق: $\sum d^2 = 10$، ويمكن كتابة ذلك في هيئة صيغة رياضية تسمي صيغة معامل سبيرمان

$$p = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

حيث n هو عدد النباتات وهو 6 في المثال السابق. (يلاحظ أن الارتباط يكون ارتباطا تاما إذا تساوت جميع الرتب وعندها يكون معامل الارتباط مساويا الواحد الصحيح، وهذا شيء مرغوب فيه). في المثال السابق يكون معامل ارتباط الرتب هو

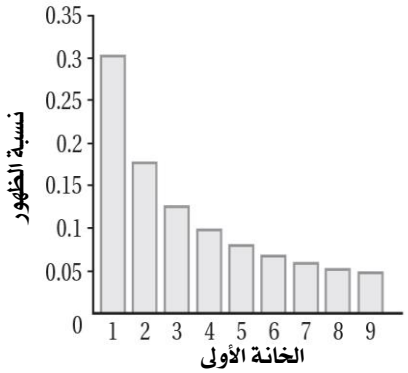
$$p = 1 - \frac{6 \times 10}{6(36 - 1)} = 0.71$$

(لأقرب رقمين عشريين) وهو ارتباط طردي متوسط.

ظاهرة الخانة الأولى من اليسار The first digit phenomenon

لنأخذ مجموعة بيانات واقعية، دفاتر حسابات شركة على سبيل المثال أو ارتفاع سلسلة جبال. ثم نسجل عدد مرات ظهور الأعداد من 1 إلى 9 في الخانة الأولى من اليسار مع إهمال الصفر. قد يتوقع معظم الناس أن التسعة أرقام لها نفس احتمالية الظهور (أي أن كل ظهور لرقم من التسعة أرقام تكون احتماليته $\frac{1}{9}$)، ولكن لاحظ عالمان من علماء القرن التاسع عشر هما:سايمون نيوكوم Simon Newcomb، وفرانك بينفورد Frank Benford أن ذلك لا يتحقق . وتقصى بينفورد الأمر بتعمق أكثر فقام بحساب الخانات الأولى من اليسار في كمية كبيرة من البيانات بدءا من إحصائيات لعبة كرة البيسبول وحتى إحصائيات مصبات الأنهار ، ووجد أن العدد 1 يظهر في الخانة الأولى بنسبة 30٪ بينما يظهر الرقم 2 بنسبة 17٪ وتقل النسبة حتى نصل إلى العدد 9 الذي يظهر بنسبة 5٪.

قانون بينفورد Benford's law



يقدم قانون بينفورد صيغة رياضية لظاهرة الخانة الأولى ، وينص على أن نسبة تكرار ظهور العدد n في الخانة الأولى من اليسار يساوي تقريبا $\log_{10}(1+\frac{1}{n})$ لكن علينا أن نتوخى الحذر عند التعامل مع هذا القانون؛ حيث أنه في الغالب لا ينطبق على الأرقام العشوائية (التي نطلق عليها البيانات الموزعة توزيعا منتظما) ولهذا يكون قانون بينفورد عديم القيمة في اختيار أرقام اليناصيب. وبالمثل لا ينطبق القانون على البيانات التي تقع في نطاق ضيق (فمثلا: العديد من رؤساء الولايات المتحدة الأمريكية لا تكون الخانة الأولى في أعمارهم هي 1) . أما مابين الحالتين السابقتين تتكرر مواقف لا حصر لها سواء أكانت اجتماعية أم طبيعية تخضع لقانون بينفورد. (لقانون بينفورد قيمة كبيرة في الولايات المتحدة في مجال التقصي عن الاحتيال).

نظرية هيل Hill's theorem

أصبح استخدام قانون بينفورد سائغ الاستخدام في نواح عدة منذ ظهوره. بالنظر إلى الأرقام من 1 إلى 9 فقط، نجد أن الخانة الأولى تكون تامة الانتظام. الآن سنكون مجموعة جديدة من الأرقام عن طريق مضاعفة الأرقام من 1 إلى 9 لتصبح المجموعة كالآتي: 2,4,6,8,10,12,14,16,18

فأصبحت احتمالية ظهور الرقم 1 في الخانة الأولى تفوق ال 50٪، وهذا يفسر عدم استقرار التوزيع المنتظم للخانات الرائدة. سوف تقوم العديد من الإجراءات الرياضية بتحريف التوزيع المنتظم إلى توزيع بينفورد.

كانت هذه محض فكرة لم تثبت قبل عام 1998م، العام الذي قام فيه العالم ثيودرهيل– Theodore Hill بتقديم تفسير دقيق. كانت الملاحظة الرئيسة أن قانون بينفورد لا يعتمد على الأساس–base. لقد بني قانون بينفورد على الأعداد التي أساسها 10، ولكن يمكن تطبيق نفس الفكرة على أي أساس آخر b مع استخدام معدل ظهور الخانة الأولى n من العلاقة $\log_b(1+\frac{1}{n})$ ، وقد أثبت هيل أن قانون بينفورد هو التوزيع الاحتمالي الوحيد للخانات الرائدة الذي يحقق خاصية عدم الاعتماد على أساس الأرقام.

الاحتمال PROBABILITY

الاحتمال Probability

الاحتمال هو طريقة رياضية للتعبير عن فرصة حدوث حدث ما وليكن X. تطبق هذه الطريقة بمساواة المتغير X بقيمة سنرمز لها بالرمز P(X) تتراوح قيمتها بين صفر، وواحد.

ببساطة، إذا كان P(X) يساوي صفرا يكون الحدث X مستحيلا،على سبيل المثال : احتمال ظهور العدد 7 عند إلقاء حجر نرد أما إذا كانت P(X) يساوي 1 يكون الحدث مؤكدا وفيما بين ذلك مثلا عند إلقاء قطعة نقود منتظمة يكون احتمال ظهور الصورة يساوي 2/ 1، ويكون احتمال ظهور رقم من 1 إلى 5 عند إلقاء حجر النرد مساويا 5/ 6.

يكون حساب الاحتمال مفيدا في نواح عديدة، عندما يكون عدد النواتج محدودا لكن في بعض التطبيقات يكون حساب الاحتمال بسيطا. من الأفضل أن نترجم العبارة P(X)=0 على أنها تعني أن الحدث X غير محتمل الحدوث أبدا. لنفرض أنني سأختار بطريقة عشوائية تماما رقما حقيقا يقع بين0 ، و 10 . ما هو احتمال اختيار القيمة الدقيقة للعدد ط (دون تقريب إلى رقم أو رقمين عشريين)؟ الاحتمال هنا صفر؛ لأن العدد ط هنا هو واحد من أرقام لا نهائية التي يمكن أن يقع على أي منها الاختيار. لدراسة احتمالات كهذه، جاء ما يسمى التوزيع الاحتمالي المتصل(continuous probability distributions).

النواتج outcomes والنواتج الناجحة successes

عند إلقاء حجر نرد منتظم نجد أن هناك ست نواتج جميعها لها نفس احتمال الظهور. لنفرض أننا مهتمين باحتمالية ظهور العدد 6 إذا يعنى أن أحد النواتج تحديدا سيكون الحصول عليه هو النتيجة الناجحة

واحتمال ظهوره=1/ 6 إن المبدأ الأساسي في حساب الاحتمال مباشر جدا؛ حيث يمكن حساب احتمال ظهور النتيجة الناجحة دائما من العلاقة احتمال النتيجة الناجحة =

$$\frac{\text{عدد النواتج الناجحة}}{\text{عدد النواتج الكلية}}$$

طالما كانت كل النواتج لها نفس احتمال الظهور.

وإذا قمنا برمي حجري نرد منتظمين (B،A) فسيكون لدينا 36 نتيجة محتملة: احتمال ظهور الأرقام من 1إلى 6 على حجر النرد A مع احتمال ظهور الأرقام من 1 إلى 6 على حجر النرد B. وإذا أردنا معرفة احتمال ظهور رقمين مجموعها 7 .

نحسب عدد حالات الظهور الفعلي لرقمين مجموعهما . 7 مرات الظهور الفعلي هي

A	1	2	3	4	5	6
B	6	5	4	3	2	1

وبما أن عددها 6،إذا الاحتمال هنا = 6/ 36=1/ 6.

يمكن استخدام هذه الفكرة البسيطة كأساس لمسائل ذات مستويات أعلى مثل: مسألة عيد الميلاد–Birthday Problem. إن حساب احتمالات كهذه ينطوي على استخدام ما يسمي بالتوافيق وذلك لحساب عدد النواتج

جمع الاحتمالات Adding Probabilities

لنفرض أننا نعرف احتمال وقوع حدثين: X، Y.فما هو احتمال وقوع الحدث الجديد X أو Y؟بحكم التجارب تم التوصل إلى أن (أو–or) تعني (جمع الاحتمالات– Adding probabilities)

على سبيل المثال: احتمال ظهور الرقم 4 عند رمي حجرة نرد منتظمة =1/ 6، وكذلك ظهور الرقم 5 له نفس الاحتمال، إذا يكون احتمال ظهور 4 أو 5 يساوي 1/ 6+1/ 6=1/ 3 ولابد من توخي الحذر الشديد في ذلك؛ لأن حدوث أخطاء وارد جدا في مثل هذه العمليات. لنفرض أنني رميت قطعتي نقود A، B، ما هو احتمال ظهور الصورة على القطعة A أو القطعة B ؟ احتمال ظهور الصورة على القطعة A= 1/ 2 ، وكذلك احتمال ظهور الصورة على القطعة B ، وبجمع الاحتمالين 1/ 2+1/ 2=1

نجد أن مجموعهما يساوي. قد يتبادر إلى الذهن أن هذا يعني أن الحدث مؤكد، ولكن في واقع الأمر، هذا ليس له معنى؛ لأن ببساطة يمكن أن أحصل على كتابتين (أي أن وقوع حدث ظهور صورتين معا ليس حدثا مؤكدا).

إذا بشكل عام لا يمكن تطبيق قاعدة جمع الاحتمالات إلا إذا كان الحدثان متنافيين.

ضرب الاحتمالات Multiplying Probabilities

لنفرض أننا نعرف احتمال وقوع حدثين: X، Y. فما هو احتمال وقوع الحدث الجديد X و Y؟ بحكم التجارب في هذه الحالة تم التوصل إلى أن (و–and) تعني ضرب الاحتمالات.

على سبيل المثال: إذا قمت بإلقاء قطعتي نقود منتظمتين فسيكون احتمال ظهور الصورة على القطعتين معا = 1/ 2 ×1/ 2=1/ 4.

وكما هو الحال في جمع الاحتمالات، يؤدي الخطأ في تطبيق القاعدة إلى الحصول على نتائج ليس لها معنى؛ فبفرض أنني ألقيت قطعة نقود منتظمة مرة واحدة، سيكون احتمال ظهور الصورة يساوي 1/ 2 وكذلك سيكون احتمال ظهور الكتابة يساوي 1/ 2، إذا بتطبيق قاعدة ضرب الاحتمالات سنجد أن احتمال ظهور الصورة والكتابة معا (في نفس الرمية) يساوي 1/ 4، وهذا غير معقول؛ لأن ذلك الحدث مستحيل الحدوث فيكون احتمال حدوثه مساويا صفر.

القواعد السابقة نتجت بحكم التجارب والخبرة إلا أننا لابد من أن نفهم معنى الأحداث المستقلة–Independent events حتى نعرف متى تكون هذه القواعد قابلة للتطبيق أم لا.

الأحداث المتنافية Mutually exclusive events

يقال أن الحدثينX وY متنافيان إذا كان احتمال حدوثهما معا مستحيلا؛ أي أنه إذا وقع الحدث X فإن الحدث Y لا يقع والعكس صحيح.فمثلا إذا رميت حجر نرد سيكون حدثا ظهور العدد 2، وظهور العدد 5 حدثين متنافيين.

وإذا كان الحدثان X وY متنافيين فإنه لحساب احتمال وقوع أحدهما (X أوY) نستخدم قاعدة جمع الاحتمالات

$$P(X \text{ OR } Y) = P(X) + P(Y)$$

على سبيل المثال: إذا كان احتمال ظهور العدد 2، واحتمال ظهور العدد 5 عند إلقاء حجر نرد يساويان 1/ 6 فإن احتمال ظهور 2 أو 5 يكون 1/ 6+1/ 6=1/ 3.

ولكن إذا ألقيت حجري نرد A، B فسيكون حدثا ظهور 2 على الحجر A، وظهور 5 على الحجرB غير متنافيين؛ حيث أنهما من الممكن أن يحدثا معا لذلك لا يمكن استخدام قاعدة جمع الاحتمالات.

الأحداث المستقلة Independent events

يقال أن الحدثين X وY مستقلان إذا كان حدوث أحدهما لا يؤثر على الآخر؛ أي أنه

سواء وقع الحدث X أم لم يقع فلن يتسبب ذلك في التأثير على احتمال الحدث Yوالعكس صحيح.

لنضرب مثالا تقليديا: إذا قمت بإلقاء حجر نرد وقطعة نقود فسيكون حدث الحصول على الرقم 6، وحدث ظهور الصورة على قطعة النقود حدثين مستقلين.

إذا كان X وY حدثين مستقلين فإنه احتمال حدوث X، و Y معا (X وY) يحسب بضرب الاحتمالين باستخدام العلاقة:

$$P(X \text{ AND } Y) = P(X) \times P(Y)$$

وبشكل خاص، يكون احتمال ظهور الصورة و العدد 6 معا =1/ 2 ×1/ 6=1/ 12.

الاستقلال لا يكون واضحا دائما. لنضرب مثالا آخر: إذا قمنا بسحب ورقة لعب من مجموعة أوراق ثم خلطنا الأوراق وقمنا بالسحب مرة أخرى، فهل تكون نتيجتي السحبتين أحداثا مستقلة ؟ تتوقف الإجابة على ما إذا كانت الورقة المسحوبة أولا قد أعيدت إلى مجموعة الأوراق قبل السحب مرة أخرى. فإذا أعيدت الورقة قبل السحب مرة أخرى يكون الحدثان مستقلين أما إذا لم تتم إعادتها قبل السحب مرة أخرى يكون احتمال الحدث الأول مؤثرا على احتمال الحدث الثاني فيكونان غير مستقلين.

ومن الصواب تقريبا القول بأنه لا يمكن أن يكون هناك حدثان مستقلان ومتنافيان في الوقت ذاته، ولكن يوجد استثناء واحد، تحديدا إذا كان أحد الحدثين مستحيلا. فحدثي ظهور الصورة عند إلقاء قطعة النقود، والحصول على رقم 7 عند رمي حجر نرد منتظم هما حدثان مستقلان بلا شك، وهما حدثان متنافيان في الوقت ذاته. بديهيا لا يمكن أن يتحقق الاثنان معا لأن حدث ظهور الرقم 7 نفسه مستحيل.

مسألة يوم الميلاد The Birthday problem

كم عدد الأفراد المتواجدين في غرفة ما بحيث يصبح احتمال وجود شخصين لهما نفس يوم الميلاد 50٪ على الأقل؟ من الملائم هنا أن نقلب السؤال رأسا على عقب فنقول احسب –لأعداد مختلفة من الأفراد– احتمال أن يكون جميع الأفراد في غرفة ما لهم يوم

ميلاد مختلف. القيمة الأولى لعدد الأفراد التي عندها يقل الاحتمال عن 50٪ هي إجابة المسألة الأصلية.

يصاحب النموذج المستخدم في الحل بعض الافتراضات الضمنية التي يجب ملاحظتها. فمن الواضح أن المسألة تهمل وجود السنة الكبيسة، وتفترض بمهارة أن جميع أيام السنة متساوية من حيث كونها أيام ولد فيها أشخاص على الرغم من أن هذا ليس صحيحا إلى حد ما في الواقع العملي.

نظرية يوم الميلاد The Birthday Theorem

لحل مسألة يوم الميلاد نفرض أولا أن هناك فردان فقط في الغرفة، وبالتالي يكون العدد الكلي للترتيبات الممكنة لأيام الميلاد=365×365. ومن ناحية أخرى إذا كان كل منهما قد ولد في يوم مختلف، فيمكن أن يكون يوم ميلاد الشخص الأول في أي يوم من أيام السنة (365 احتمال)، وأن يكون يوم ميلاد الشخص الثاني في أي يوم ماعدا اليوم الذي يوافق يوم ميلاد الشخص الأول (364 احتمال)

وبذلك يمكن القول بأن عدد الأيام الممكنة لأيام الميلاد حيث لا يشترك الشخصان في يوم ميلاد واحد=$\frac{365\times364}{365\times365}$ 365×364. واحتمالية حدوث أحدها يساوي

ويمكن تعميم ذلك على عدد n من الأفراد في الغرفة فيكون العدد الكلي للترتيبات الممكنة لأيام الميلاد= n365، وإذا كان لكل شخص في الغرفة يوم ميلاد مختلف فإننا سوف نتبع نفس التفكير السابق:

يمكن أن يكون يوم ميلاد الشخص الأول في أي يوم من أيام السنة (365 احتمال)، وأن يكون يوم ميلاد الشخص الثاني في أي يوم ماعدا اليوم الذي يوافق يوم ميلاد الشخص الأول (364 احتمال) ، وأن يكون يوم ميلاد الشخص الثالث في أي يوم ماعدا يومي ميلاد الشخصين الأول والثاني (363 احتمال)، وهكذا حتى نصل إلى الشخص n الذي لابد ألا يوافق يوم مولده أيام أول (n-1) أي (366−n احتمال)، واحتمال حدوث ذلك.

$$\frac{365 \times 364 \times \ldots \times (366 - n)}{365}$$

ويمكن تبسيط ذلك إلى

$$\frac{364!}{(365 - n)! \times 365}$$

إذا ما هي أقل قيمة لـ n حتى تصبح قيمة العلاقة السابقة 0.5 ؟ بتجربة عدة قيم لـ n نجد أن إجابة السؤال السابق هي n=23 وهي القيمة التي تصبح عندها قيمة الاحتمال السابق مساوية 0.493.

الاحتمالات الشرطية Conditional Probability

في مدينة ما، 48٪ من المنازل لديها اتصال انترنت عالي السرعة- broadband internet، و 6٪ من المنازل لديها كابل تليفزيون واتصال انترنت عالي السرعة، والسؤال هو : ما احتمالية أن يكون لمنزل ما كابل تليفزيون مع العلم أنه لديه انترنت عالي السرعة ؟ إذا كان X، وY حدثان فإنه يمكن كتابة الاحتمال الشرطي لكل منهما رياضيا على الصورة

$$P(X \mid Y) = \frac{P(X\&Y)}{P(Y)}$$

حيث $P(Y) \neq 0$

في المثال السابق: نفرض أن X هو حدث أن المنزل فيه كابل تليفزيون، و Y هو حدث أن المنزل فيه إنترنت عالي السرعة مع ملاحظة أننا لسنا في حاجة إلى معرفة P(X) لحساب الإجابة

$$P(X \mid Y) = \frac{0.06}{0.48} = 0.125, \text{OR } 12.5\%$$

وفي العديد من المواقف يكون حساب الاحتمال الشرطي مفيدا للغاية لأنه يسمح بتحديث الاحتمالات عند توفر معلومات جديدة ، ويطلق على ذلك الاستدلال البايزي- Bayesian inference.

نظرية بايز Bayes' Theorem

عام 1764م نشر بحث مهم للمبجل توماس بايز(Thomas Bayes) بعد وفاته. وقد ذكر فيها تفسير قاطع للاحتمالات الشرطية. وكان ذلك أساس نظرية بايز التي تنص على أنه: لأي حدثين X و Y يمكن تطبيق العلاقة

$$P(X \mid Y) = P(Y \mid X) \times \frac{P(X)}{P(Y)}$$

إلى حد ما، هذه العلاقة ليست معقدة، إنها تتبع تعريف الاحتمال الشرطي:

$$P(Y \mid X) = \frac{P(X\&Y)}{P(X)}$$

وكذلك

$$P(X\&Y) = P(Y \mid X)P(X)$$

وبالتعويض بذلك في العلاقة $P(X \mid Y)$ نحصل على النتيجة. وهذه النظرية لها أهمية بالغة، على سبيل المثال، تستخدم في تحليل مسألة الايجابيات الزائفة problem of false positives.

تجزئة الحدث Splitting an event

لنفرض أننا اخترنا شخصا عشوائيا في شارع وأردنا أن نحسب احتمال أن هذا الشخص يرتدي نظارات. والإحصائيات المتوفرة في مدينتنا تقول أن 65٪ من الإناث ، و40٪ من الذكور يرتدون النظارات، ونعلم أيضاً أن51٪ من السكان إناث، و49٪ ذكور،كيف يمكننا دمج المعلومات التي لدينا حتى نحسب الاحتمال المطلوب؟

لنفرض أن Y هو حدث أن الشخص أنثى. و X هو الحدث محل اهتمامنا؛حدث أن الشخص يرتدي نظارات. لقد قمنا بتقسيم الأحداث إلى حدثين أصغر هما (X و Y معا أي X & Y) ، و (X وليس Yأي X & not) وهما حدثان متنافيان إذا:

$$P(X) = P(X\&Y) + P(X\&\text{not } Y)$$

وبالتعبير عن $P(X\&Y)$ باستخدام الاحتمال الشرطي نحصل على:

$$P(X\&Y) = P(X \mid Y)P(Y)$$

وبالمثل

$$P(X\&Y) = P(X \mid Y)P(Y)$$

وبوضع ذلك في الصيغة السابق ذكرها نحصل على:

$$P(X) = P(X \mid Y)P(Y) + P(X \mid \text{not } Y)P(\text{not } Y)$$

عندئذٍ يمكننا استخدام البيانات المتوفرة لدينا: $P(X|Y)=.65$، $P(X|notY)=0.4$

$$P(Y)=0.51 \text{ ، } P(\text{not } Y)=0.49$$

فينتج

$$P(X)=0.65\times0.51+0.4\times0.49=0.5275$$

الإيجابيات الزائفة False positives

إذا كانت مؤشرات فحص مرض ما كالتالي: إذا كان الشخص مريضا يعطي الاختبار 99٪ من الوقت نتيجة إيجابية ، ونتيجة سلبية زائفة 1٪، أما إذا كان الشخص سليما يعطي الاختبار نتيجة سلبية 95٪ من الوقت و نتيجة إيجابية زائفة 5٪. وهذا المرض مرض نادر يصيب نسبة 0.03٪ من السكان. فإذا اخترنا شخصا اختيارا عشوائيا وليكن اسمه هارولد وقمنا بفحصه ، وكانت نتيجة فحصه إيجابية فيكون السؤال المحير هنا: ما هو احتمال إصابته بالمرض؟

نريد أن نضع ذلك في صورة رياضية: لنفرض X هو حدث أن هارولد مصاب. قيمة P(X) قبل أن نقوم بتحليل بيانات الفحص = 0.0003

Y هو حدث أن فحص هارولد إيجابي، ويمكننا تجزئة هذا الاحتمال وكتابته على الصورة

$$P(Y) = P(Y \mid X)P(X) + P(\text{not } X)$$

وناتجه هو $0.050282=0.9997\times0.05+0.0003\times0.99$

يهمنا الآن حساب P(X|Y) وهو احتمال إصابته بالمرض مع كون نتيجة فحصه

موجبة، وهنا يمكن تطبيق نظرية بايز لنحصل على النتيجة الآتية:

$$P(X \mid Y) = P(Y \mid X) \times \frac{P(X)}{P(Y)} = 0.99 \times \frac{0.0003}{0.050282} = 0.50091$$

إذا احتمال إصابة هارولد بالمرض علما بأن نتائج فحصه إيجابية أقل من 0.6٪.

وتفسير هذه النتيجة المثيرة للدهشة هو أن النتائج الإيجابية الحقيقية تمثل نسبة كبيرة من عدد المصابين الضئيل الذي يفوقه بنسبة كبيرة النتائج الإيجابية الخادعة ، وهذه النتائج الإيجابية الخادعة تشكل نسبة صغيرة من عدد ضخم هو عدد غير المصابين لذلك ،على الرغم من دقة الفحص ، فعندما تم اختيار عينات عشوائية من الناس كانت غالبية النتائج الإيجابية خادعة.

مغالطة المدعي العام The prosecutor's fallacy

نفرض أن لدينا متهم تجري محاكمته بتهمة السطو. وجدت الشرطة في مسرح الجريمة خصلات من شعر اللص، وأظهرت فحوصات الطب الشرعي أنها تطابق شعر المتهم، وشهد علماء الطب الشرعي أن احتمال وجود شخص عشوائي يحقق نفس التطابق مع خصلة الشعر يساوي $\frac{1}{2000}$.

نقع في مغالطة المدعي عندما نستنتج أن احتمال كون المتهم مذنبا يساوي $\frac{1999}{2000}$ واعتبار ذلك دليلا دامغا.

بالطبع، هذا غير صحيح. فإذا كان عدد سكان مدينة يساوي 6 مليون نسمة، سيكون عدد الأشخاص الذين قد تطابق خصلة الشعر في مسرح الجريمة مع العينة المأخوذة منهم يساوي

$$3000 = 6{,}000{,}000 \times \frac{1}{2000}$$

إذا يكون احتمال أن يكون هذا المتهم مذنبا استنادا فقط على هذا الدليل $=\frac{1}{3000}$.

صاغ مصطلح (مغالطة المدعي–The prosecutor's fallacy) كل من ويليام تومسون–William Thomson، وإدوارد شومان–Edward Schumann عام 1987م في

مقالهما الذي كان عنوانه"تأويل الأدلة الإحصائية في القضايا الجنائية Interpretation of-statistical Evidence in Criminal Trials"التي تناولوا فيها مدى سهولة وقوع الناس في هذه المغالطة بما فيهم على أقل تقدير مدع عام واحد.

مغالطة محامي الدفاع The defence attorney's fallacy

وضع تومسون، وشومان في الاعتبار أيضاً الخطأ العكسي لمغالطة المدعي العام وأطلقوا عليها اسم مغالطة محامي الدفاع. في المثال السابق ذكره ، قد يجادل محامي الدفاع عن أن دليل تطابق خصلة الشعر دليل عديم القيمة؛ لأنه يرفع من احتمال الإدانة بنسبة ضئيلة $\frac{1}{3000}$. إذا كانت خصلة الشعر هي الدليل الوحيد ضد المتهم، يصبح عدد المشتبه بهم قبل أخذ دليل خصلة الشعر في الاعتبار هو عدد سكان المدينة جميعا 6000000، وبأخذ الدليل الجديد في الاعتبار يقل هذا العدد بمعامل 3000 ليصبح 2000.

ومع ذلك، قد يتوقع أحدهم أن هذا ليس هو الدليل الوحيد، وفي هذه الحالة يصبح العدد الابتدائي للمشتبه بهم الممكنين أصغر بكثير، فإذا كان يساوي 4000 على سبيل المثال قد يقلله دليل الطب الشرعي بمعامل 2000 فيصبح 2، مما يرفع احتمال أن يكون المتهم مذنب من $\frac{1}{4000}$ إلى $\frac{1}{2}$ فيصبح دليلا ذا قيمة.

مغالطة مقلوب الاحتمال The fallacy of probability inversion

قد تثير ظاهرة الإيجابيات الخادعة ومغالطة المدعي الدهشة. ومن الناحية الرياضية نجد أنهما متشابهتان: حيث أن المغالطة في الحالتين كانت الخلط بين$P(X|Y)$،و$P(Y|X)$. عندما تكون قيمة $P(X|Y)$ كبيرة جدا يفترض الناس أن $P(Y|X)$ لابد أن تكون قيمتها كبيرة أيضاً. في مثال الفحص الطبي الذي ذكر آنفا كان احتمال وقوع نتائج إيجابية بشرط وقوع الإصابة = 0.99، بينما احتمال وقوع الإصابة بشرط وقوع نتائج إيجابية= 0.056. وهذه الأمثلة توضح الخلط الذي قد يحدث.

وهذه المغالطة واسعة الانتشار (حتى في غرف الجراحة وقاعات المحاكم). ويجادل بعض الناس عن أنها ليست مجرد خطأ رياضي شائع، بل تحيز معرفي متأصل في الطبيعة

البشرية. في كلتا الحالتين يظل تقدير هذا الأمر أساسيا لإعطاء البيانات الإحصائية معنى في العالم الواقعي.

التكرارية Frequentism

ما الحالة الوجودية للاحتمال؟ بمعني آخر، إلى مدي يوجد حقا ما يسمى بالاحتمال في الواقع؟ هناك مدرستان فكريتان رئيسيتان هما: المدرسة التكرارية Frequentism، والمدرسة البيزية Bayesianis .

يرى مؤيدو المدرسة التكرارية أن العشوائية جزء أصيل في العالم الواقعي، وهذا ما يحدده الاحتمال. فعندما نقول أن احتمال الحدث A يساوي $\frac{1}{2}$ يعنى أنه إذا أعيدت التجربة عدة مرات فإن A يقع نصف عدد المرات بالضبط، ما يعني أن احتمال الحدث A هو مقياس تكرار وقوعه مع معرفة الشروط الابتدائية (قد تكون هذه القيمة قيمة تقريبية تم التوصل إليها بعد عدد محدد من التجارب، و قد تكون قيمة مضبوطة إلى أقصى حد)

وكما هو واضح، لا يطبق هذا المبدأ بسهولة على الأحداث التي تقع مرة واحدة فهو مناسب أكثر للأحداث المتكررة

البيزية Bayesianism

أما مؤيدو المدرسة البيزية – على النقيض من مؤيدي المدرسة التكرارية– يرون أن الاحتمال غير موجود في العالم الخارجي. إنه – بكل ما تعنيه الكلمة– وسيلة بشرية لتحديد درجة التأكد وذلك استنادا على بيانات تفتقر إلى الكمال. بمعنى آخر: الاحتمال هو مفهوم موضوعي؛ حيث تكون تقديرات الناس له مختلفة تبعا لاختلاف المعلومات المتوفرة لديهم.

الاحتمال المبدئي لرمية قطعة نقود وظهور الصورة عليها=1/ 2 نتج عن أن المعلومات المتوفرة لدينا ضئيلة، وإذا توفرت لدينا معلومات أخرى أكثر عن وزن قطعة النقود، وموضعها الابتدائي، وأسلوب الرامي يكون بإمكاننا تعديل الاحتمال، وإذا عرفنا قدر كبير من تلك المعلومات يصبح بإمكاننا التنبؤ بالنتيجة بشيء من التأكيد. (عرف عن عالم

الرياضيات جون كونواي–John Conway إتقانه مهارة إلقاء قطعة النقود والحصول على نتيجة بعينها).

الاستدلال البايزي Bayesian inference

نتيجة لوجهة النظر البايزية يري مؤيدو هذه النظرية أن كل الاحتمال يندرج تحت نوعية الاحتمال الشرطي، لنفرض أننا أشرنا إلى احتمال وقوع الحدث A ب P(A) (يكون ذلك في حقيقة الأمر P(A|C) حيث C يمثل المعرفة الحالية ولكننا نهمل كتابتها). ويكون هذا هو الاحتمال المسبق، وعندما تردنا معلومات جديدة ولتكن B سنحتاج إلى تحديث حساباتنا وهذا يعني أنه عند استخدام الاحتمال الشرطي لحساب P(A|) يسمي ذلك بالاحتمال اللاحق.

وكما يظهر من مغالطة مقلوب الاحتمال، يمكن للاستدلال البايزي أن يهمل النتائج غير المتفقة مع الحدس، يستخدم المفكرون المؤيدون لوجهة النظر البايزي هذا الأسلوب لتحسين تقدير الاحتمال في مجموعة كبيرة من الموضوعات بدءا من الاقتصاد وحتى الذكاء الاصطناعي.

نظرية الاتفاق لأومان Aumann's agreement theorem

هناك ثلاثة أجزاء في الاستدلال البايزي: التوزيع الاحتمالي السابق، وبيانات جديدة، والتوزيع الاحتمالي اللاحق الذي هو نتاج الجزأين السابقين.

في عام 1976م، قام أومان بتزويد اثنين من المفكرين المؤيدين للنظرية البيزية باحتمالات مسبقة متطابقة لحدث ما X، وزود كل شخص بمجموعة بيانات مختلفة. الاحتمال الشرطي على الأرجح سيعطي احتمالين لاحقين مختلفين للحدث X بالطبع. كان سؤال عن ما عما سيحدث إذا قام الشخصان بمشاركة احتماليهما اللاحقين (بمعني جعلهم على نفس القدر من المعرفة العامة) بدون مشاركتهم لمجموعة البيانات المعطاه لكل منهما.

الإجابة مباشرة رياضيا، ولكن بأي حال من الأحوال: بعد مرات عديدة من مشاركة الاحتمالات وإعادة الحسابات سيصل الاثنان في النهاية إلى نفس احتمال الحدث X اللاحق. فقال أومان العبارة الآتية: "ليس باستطاعة من لهم نفس السوابق أن يتفقوا على ألا يتفقوا".

مسألة مونتي هول Monty Hall Problem

مونتي هول هو المقدم السابق للمسابقة التليفزيونية لنعقد صفقة!– Let's make a deal. كان على المنافسين اختيار باب واحد من بين ثلاثة أبواب ليظهر ما خفي وراءه من جوائز مختلفة. كانت تلك هي سلسلة الأحداث المتبعة التي شكلت الإلهام لبرنامج مسابقات مونتي هول التي اشتهرت بأنها أسوأ ألغاز الاحتمال سمعة.، وكان مكتشفه عام 1975م هو ستيف سليفن–Steve Slevin.

هناك ثلاثة أبواب A، وB، وC، خلف أحدهم سيارة رياضية جديدة كليا أما الاثنان الآخران فخلفهما ملعقتين خشبيتين. يختار المتسابق أحد الأبواب وليكنA بذلك تكون احتمالية اختياره الجائزة الكبرى=1/ 3.

ثم يقول مونتي –الذي يعلم الباب الذي توجد خلفه السيارة–: "لن أخبرك عما يوجد خلف الباب A، ولكن سأكشف لك عن أن وراء الباب B ملعقة خشبية، هل ستحتفظ باختيارك أم تريد اختيار الباب C ؟"

الفرض الطبيعي هنا هو أن الاحتمالات بين كل من A، وC له متساوية، والتبديل بينهما لن يشكل أي فرق. وفي الواقع هذا الفرض غير صحيح: حيث أن احتمال أن يكون C هو الباب الذي يخفي السيارة =2/ 3 بينما احتمال A =1/ 3، لذلك ينبغي على المتسابق أن يبدل اختياره.

هل ينبغي تبديل الاختيار أم لا؟

القول بأن حل مسألة مونتي هول كان مفاجأة فيه شيء من التهوين. شعر قراء مجلة باراد – Parade بالغضب عندما ناقشت مارلين فو ساف هذه المسألة عام 1990م. وانهالت الشكاوى التي صدر بعضها عن العديد من علماء الرياضيات المحترفين اعتراضا منهم على حل المسألة الذي وصفوه بأنه خاطئ ويدل على جهلها بعلم الرياضيات.

من الأفضل زيادة عدد الأبواب لإثبات صحة ما قالته مارلين وليكن عددهم 100 باب. لنفرض أن المتسابق عليه اختيار الباب رقم 54 واحتمال إيجاد السيارة هو 1٪ . بعد

ذلك يكشف مونتي أن الأبواب التي لها الأرقام: من 1 إلى 53، ومن 55 إلى 86، ومن 88 إلى 100 تخفي وراءها ملعقة خشبية. هل على المتسابق تبديل الباب 54 بالباب 87 أم التمسك بالباب 54؟ المفتاح الرئيسي للحل هنا هو أن احتمال أن يكون الباب 45 هو الباب الرابح يظل كما هو 1٪؛ حيث أن مونتي كان حريصا على ألا يشي بما يؤثر على هذا الاحتمال.أما نسبة ال 99٪ الباقية فقد أصبحت من نصيب الباب 87 بدلا من أن تتوزع على بقية الأبواب الأخرى، وبذلك يتضح أنه ينبغي التبديل واختيار الباب 87.

تعتمد مسألة مونتي هول على الحذاقة. النقطة الحاسمة هي أن يكون مونتي هول على علم بمكان وجود السيارة. إذا كان مونتي لا يعرف ذلك وفتح أحد الأبواب الأخرى عشوائيا (مخاطرا بالباب الصحيح ولكن وجد باب وراءه ملعقة خشبية) سيتغير الاحتمال ويصبح مساويا 1/ 2، أما في المسألة الأصلية يفتح أي باب من الاثنين الذي يعلم أن وراءهما ملعقة خشبية ويظل احتمال إيجاد المتسابق للسيارة مساويا 1/ 3 كما كان في البداية. دون أيما تغيير.

إبرة الكونت بوفون Count Buffon's needle

إذا قمت بإلقاء إبرة عشوائيا على ورقة مسطرة، فما هو احتمال أن تسقط الإبرة في المسافة الفاصلة بين أي خطين؟ لقد بحث جورج لوكلير(كونت دي بوفون)– Georges Leclerc le Comte de Buffon هذا السؤال في عام 1777م. أجرى بوفون تجاربه بإلقاء عصي فوق كتفيه، وتركها تسقط على بلاط أرضية غرفته.

تعتمد الإجابة على طول الإبرة (L)، والمسافة (d) التي تفصل بين الخطوط.

إذا كانت L≤d تكون الإجابة $\frac{2l}{\pi d}$.

أما في حالة L > d يصبح الأمر معقدا أكثر ولكن بدرجة طفيفة. بالتالي إذا كان طول الإبرة 1سم، والمسافة بين الخطين 2 سم تكون الإجابة ببساطة 1/ π، وهذا يمثل الطريقة المعروفة باسم طريقة مونت كارلو– لحساب قيمة π: وهي أن تقوم بإجراء التجربة عدد ما يحلو لك من المرات ثم تقوم بقسمة عدد المرات الكلية للسقوط على عدد مرات سقوط

الإبرة على الخط وبذلك تحصل على قيمة تقريبية لـ π. وحاول كومت دي بوفون في العام نفسه تطبيق الاحتمال الشرطي على دراسة الفلسفة عن طريق محاول حساب احتمال أن تشرق الشمس صباح اليوم التالي علما بأنها قد أشرقت عدد n من المرات قبل ذلك.

قانون الأعداد الكبيرة The law of large numbers

تمت تسمية هذا المبدأ لأول مرة عام 1989 م من قبل كل من بيرسي ديوكوينز Persi Dioaconis وفريدريك موستلير Frederick Mosteller على الرغم من أن هذه الظاهرة كانت معروفة لفترة كبيرة قبل ذلك الوقت.

عندما تكون العينة كبيرة بما فيه الكفاية، يصبح أي شيء صادم محتمل الحدوث

من المعتاد أن اليناصيب لا تحظي بنسب احتمالات جيدة؛ حيث أنني إذا قمت بشراء بطاقة يناصيب فستكون فرصتي في الفوز ضئيلة جدا (ربما 1 إلى 10 ملايين). إذا فزت سأصاب بالدهشة من وقوع حدث احتمالية حدوثه ضعيفة، ولكن من وجهة اليناصيب: إذا بيعت عدة ملايين بطاقة يناصيب فإن احتمالية فوز أحد الأشخاص احتمالية جيدة.

أما المثال الأكثر غرابة وشذوذا هو فوز امرأة أمريكية بيناصيب نيوجيرسي مرتين. من وجهة نظر الفائز هذا حقا حدث لا يصدق. وضع كل من دياكوينز (Diaconis)، وموستيلر (Mosteller) سؤالا أعمق: "ما هي فرصة أن يفوز شخص واحد من بين ملايين وملايين الأشخاص الذين يشترون بطاقات اليناصيب مرتين في عمره؟"

كانت إجابة كل من ستيفن صمويل (Stephen Samuels) وجورج مكاب (George McCabe) في فترة تالية لذلك: "من الناحية العملية، هذا مؤكد".

المصادفة Coincidence

يمثل قانون الأعداد الكبيرة أداة مفيدة لدراسة المصادفة عند وقوع أحداث كان من المرجح بدرجة كبيرة أنها لن تحدث. لنفرض أننا نطلق على حدث ما أنه حدث نادر إذا كان احتمال حدوثه أقل من واحد في المليون. في عام 1953م لاحظ ليتلوود Littlewood

أن من بين سكان الولايات المتحدة الأمريكية ال 250 مليون هناك مئات الأشخاص يتعرضون لأحداث نادرة كل يوم. وبقياس ذلك على بقية العالم، نجد أنه يجب توقع حدوث الأحداث التي يكون احتمال حدوثها واحد في المليار.

مصادر مصادفة أخرى واضحة تنشأ من حدسنا الضعيف للاحتمال، وتبين الحلول المدهشة لمسألتي عيد الميلاد وظاهرة الايجابيات الزائفة كيف أن حدسنا لا يمكن الاعتماد عليه.

توزيعات الاحتمال PROBABILITY DISTRIBUTIONS

المتغيرات العشوائية Random variables

يحسب الاحتمال البسيط عن طريق معرفة عدد المرات الناجحة وعدد الاحتمالات الكلية الممكنة لوقوع حدث معين وهو وسيلة ذات قيمة إلا أنه يوجد حالات أكثر تعقيدا لن تجدي معها هذه الطريقة. إذا استبدلنا حجر النرد المنتظم بآخر غير منتظم فإن احتمال ظهور العدد 6 لن يكون $\frac{1}{6}$.

ولتنقيح ذلك نحتاج إلى مصطلحات جديدة. فضاء العينة the sample space هو مجموعة كل النتائج الممكنة لتجربة عشوائية ما. في حالة إلقاء حجر نرد يكون فضاء العينة هو الأرقام من 1 إلى 6.

المتغير العشوائي (Random variable) قيمة من القيم الممكنة المختلفة التي لكل واحدة منهن احتمال ما. (وهو من الناحية الفنية دالة من فضاء العينة للأرقام الواقعة بين الصفر والواحد). توجد كل أنواع الدوال الممكنة. المتغير العشوائي الذي يعطي للرقم 6 احتمال الظهور مرة واحدة، ويعطي الأرقام من صفر إلى 5 احتمال ظهور صفر من المرات يكافئ حجر نرد يظهر رقم 6 كحدث مؤكد.

المتغير العشوائي الذي يعطي كل رقم احتمالا يساوي $\frac{1}{6}$ يكافئ حجر نرد منتظم. هذا المتغير العشوائي له توزيعات احتمال مختلفة.

توزيعات الاحتمال Probability distributions

هناك مجموعة كبيرة من توزيعات الاحتمال، ولكنها تنقسم إلى نوعين أساسيين مختلفين. التوزيع الاحتمالي المنفصل (discrete distribution) تكون نواتج التجربة منفصلة عن بعضها كما في حالة حجر النرد الذي يمكن أن يأخذ القيمة 4، أو 5 ولكن لا يمكن أن يأخذ القيمة $4\frac{1}{2}$ في التوزيع الاحتمالي المتصل(Continuous distribution) لا يكون الأمر كذلك (على سبيل المثال: طول شخص ما يمكن أن يأخذ أي قيمة بين4، و5).

إذا كان لدينا توزيع احتمالي X، فإن أهم معلومتين لابد من معرفتهما : مركز التوزيع، ومدى انتشاره، ويتم قياسهما عن طريق: المتوسط(the mean)، والتباين (variance) على الترتيب. ترمز E إلى التوقع أو القيمة المتوقعة للتوزيع، وهو بديل (مضلل قليلا) عن المتوسط (mean).

في التجارب،يتطابق كل من المتوسط وتباين العينة لمجموعة ما من البيانات مع المتوسط النظري والتباين للتوزيع الضمني، يحدث هذا بشكل أكبر كلما زادت مجموعات البيانات المستخدمة، وذلك نتيجة لقانون الأرقام الكبيرة–law of large numbers.

التوقع Expectation و التباين Variance

إذا كان لدينا متغير عشوائي منفصل X، فإن المتوسط يعرف بأنه مجموع النواتج الممكنة كل مضروب في الاحتمال المناظر له. إذا كانتX تمثل توزيع رمي حجر نرد منتظم، النواتج الممكنة هي الأرقام من 1 إلى 6، وكل منها له احتمال $\frac{1}{6}$ إذا المتوسط

$$3\frac{1}{2} = \frac{1}{6}\times 6 + \frac{1}{6}\times 5 + \frac{1}{6}\times 4 + \frac{1}{6}\times 3 + \frac{1}{6}\times 2 + \frac{1}{6}\times 1 = 3\frac{1}{2}$$

التعريف الاصطلاحي للمتوسط(التوقع) هو:$E(X) = \Sigma x \times P(X = x)$؛ حيث X تمثل أى ناتج من النواتج الممكنة، و P(X=x) هو الاحتمال المناظر.

وإذا كان التوزيع العشوائي متصلاً فإن التعبير السابق يصبح $E(x) = \int x \int (x)dx$؛ حيث f هي دالة كثافة الاحتمال (انظر التوزيع الاحتمالي المتصل).

لنفرض أن $E(X) = \mu$ فيكون التباين $V(X)$ الذي يقيس تشتت التوزيع الاحتمالي معرفا كالتالي في رمية حجر النرد $V(X) = E(X - \mu)^2$ ويمكن حسابه بطريقة أسهل باستخدام $V(X) = E(X^2) - E(X)^2$ في رمية حجر النرد

$$E(X^2) = \frac{1}{6} \times 1^2 + \frac{1}{6} \times 2^2 + \frac{1}{6} \times 3^2 + \frac{1}{6} \times 4^2 + \frac{1}{6} \times 5^2 + \frac{1}{6} \times 6^2 = 15\frac{1}{6}$$

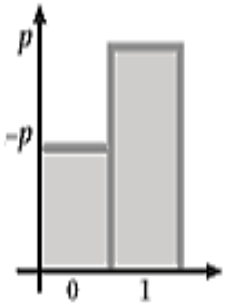
إذًا:

$$V(X) = 15\frac{1}{6} - \left(3\frac{1}{2}\right)^2 = 2\frac{11}{12}$$

يحسب التشتت كذلك باستخدام الانحراف المعياري (سيجما)الذي يعرف بالعلاقة $\sigma = \sqrt{V(X)}$

في رمية حجر النرد تكون قيمة التشتت هي $\sqrt{2\frac{11}{12}}$ ويساوي تقريبًا 1.7.

توزيع بيرنولي Bernolli distribution



هو واحد من أبسط التوزيعات الاحتمالية. يمكنه وصف تجربة إلقاء قطعة نقود غير منتظمة أو أي تجربة لها اثنان من النواتج الممكنة ذات الاحتمالية المختلفة. تجربة تسمى تجربة برنولي، لها اثنان من النواتج الممكنة: النجاح واحتماله P، والفشل واحتماله 1-P من الشائع إعطاء النجاح رقم 1 والفشل رقم صفر، وبالتالي يعطي توزيع بيرنولي احتمال 1 لـ P، و 0 لـ 1-P (وكل الأرقام الأخرى تعطى احتمال صفر).

لا يبدو توزيع بيرنولي لتجربة واحدة مفيدا كثيرا أما عند دمج عدة تجارب معا نحصل على توزيع أرقى وأعقد، ومن الأمثلة المهمة على ذلك: توزيع الدالة ذات الحدين– Binomial. التوقع لتوزيع بيرنولي هو p وتباينه P (1-P) .

تجارب ذات الحدين Binomial trials

لنفرض أننا قمنا بإلقاء 100 حجر نرد منتظم. ما هو احتمال ظهور الرقم 6 على 17 حجر منهم بالضبط؟ لأن أحجار النرد منتظمة يمكن حل المسألة بحساب عدد المرات الناجحة وعدد النواتج الكلية كما يلي. إذا نظرنا إلى ال 17 حجر نجد أن احتمال ظهور الرقم 6 عليها هو $\left(\frac{1}{6}\right)^{17}$ (حيث أن الرميات مستقلة) . ونريد أيضاً ألا يظهر الرقم 6 على ال 83 حجر المتبقي: وهذا الاحتمال = $\left(\frac{5}{6}\right)^{83}$.

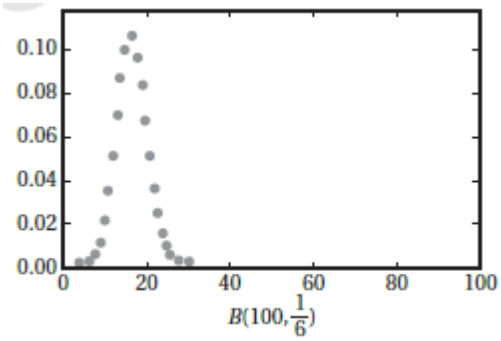
بالتالي يكون احتمال أن ال 17 حجر الذين سبق أن حددناهم-هم فقط وليس غيرهم- يظهر عليهم الرقم 6 يساوي $\left(\frac{1}{6}\right)^{17} \times \left(\frac{5}{6}\right)^{83}$. هذا الاحتمال لكل اختيار ممكن لل17 حجر. إذًا لإجابة السؤال الأصلي سوف نضرب هذا الاحتمال في عدد الاختيارات الممكنة لل 17 حجر من ال100(حيث أن الأحداث أحداث متنافية) يحسب عدد الاختيارات الممكنة لل 17 حجر من 100 باستخدام التوافيق-Combination $\binom{100}{17}$

$$\binom{100}{17} \times \binom{1}{6}^{17} \times \binom{5}{6}^{83}$$

وبالتالي تكون الإجابة : (تساوي تقريبا 0.1).

كان هذا مثالا على تجارب ذات الحدين ويتم توصيف ذلك بتوزيع ذات الحدين.

توزيع ذات الحدين Binomial distribution



يحدد توزيع ذات الحدين بقيمتين: n وهو عدد تجارب بيرنولي المتطابقة التي أجريت، P احتمال أن تنتهي كل من هذه التجارب بنجاح. في المثال السابق $P = \frac{1}{6}, n = 100$، ونكتب التعبير $X \sim B(100, \frac{1}{6})$ لنعني أن X متغير عشوائي مع توزيع ذات الحدين

وبشكل عام إذا كان $X \sim B(n, p)$، إذًا X تقيس عدد المرات الكلي للنجاح من بين n من المرات المستقلة من تجارب بيرنولي التي لها احتمالP إذًا لكل عدد K يقع بين صفر و n يكون احتمال أن X=K هي

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

أما توقع X فيحسب من العلاقة $E(X) = np$، والتباين: $V(X) = np(1-p)$

عملية بواسون Poisson Distribution

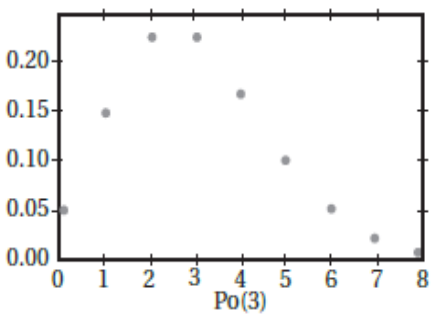
يستقبل تليفون مكتب متوسط مكالمات يقدر بـ 3 مكالمات في الساعة في الفترة بين التاسعة صباحا وحتى الخامسة مساء. في ساعة ما قد لا يتلقى اتصالات، أو يتلقي اتصالا واحدا أو 2، 3،4،5، 6 ..... مكالمات، ولا نعرف تحديدا آخر رقم لعدد المكالمات يجب أن نتوقف عنده ولكن كلما زاد العدد قل الاحتمال تدريجيا، فإذا افترضنا أن الاتصالات تأتي في أوقات عشوائية بدون فارق بين أوقات تلقي المكالمات على مدار اليوم الواحد أو بين الأيام المختلفة يكون ذلك مثالا على عملية بواسون.

بشكل عام: تضع عملية بواسون نموذجا لعدد مرات حدوث ظواهر عشوائية خلال وقت محدد معطى أو خلال منطقة ما في الفضاء.

توزيع بواسون Poisson-Distribution

نحتاج متغيرا عشوائيا يحدد احتمال جميع مرات الظهور الممكنة 0,1,2,3,4, لعملية بواسون. التوزيع الأكثر شيوعا في الاستخدام هو توزيع بواسون الذي اكتشفه العالم عام 1838م سيمون– دينيس بواسسون باستخدام قانون الأحداث النادرة – the law of rare events.

نحتاج رقما واحدا لتحديد توزيع بواسون يطلق عليه عادة الكثافة (Intensity) λ. في المثال السابق: $\lambda = 3$، ولنعبر عن أن X متغير عشوائي في هذا التوزيع نكتب $X \sim PO(3)$.



بالتالي يكون احتمال تلقي عد صفر من المكالمات في ساعة ما $P(X=0)=e^{-3}$ واحتمال تلقي مكالمة واحدة $P(X=1)=3e^{-3}$، ومكالمتين $P(X=2)=\frac{3^2e^{-3}}{2}$ واحتمال تلقي K من المكالمات $P(X=K)\frac{3^ke^{-3}}{k!}$.

وبشكل عام إذا كان $\text{if } X\sim po(\lambda)$ ، فإن

$$P(X-k)=\frac{\lambda^k e^{-\lambda}}{k!}$$

التوقع والتباين لتوزيع بواسون بالمعامل لامدا يعطى بالعلاقة $E(X)=V(X)=\lambda$.

قانون الأحداث النادرة Law of rare events

لنفرض أن مصنعًا ما ينتج آيس كريم بطعم الفول السوداني. متوسط عدد السوداني في بولة الآيس كريم=18 . والطريقة الطبيعية لنمذجة هذه المسألة تكون باستخدام توزيع بواسون $X\sim po(18)$.

لكن هناك منهج آخر . لنفرض أن هناك 36000 حبة فول سوداني في الوعاء الذي يصنع فيه الآيس كريم، واحتمال وجود حبة ما في بولة ما يساوي 0.0005. يمكننا نمذجة هذا السيناريو باستخدام توزيع ذات الحدين $X\sim B(36,000,0.0005)$.

ومن المؤكد أن النموذج الأول أكثر ملائمة لكن النموذج الثاني أكثر دقة، وهذا لا يشكل أمرًا مقلقًا؛ لأن قانون الأحداث النادرة يضمن أن نتيجتي النموذجين متقاربتين إلى حد كبير.حيث أنه من الناحية العلمية ينص قانون الأحداث النادرة على أنه عندما تزدادn كثيرًا وتصبح p صغيرة بحيث يظل متوسط عدد المرات الناجحة ثابتًا $\lambda = np$ ، وبالتالي يزداد التوزيع $B(n,p)$ ليصل إلى توزيع بواسون $po(\lambda)$، وكانت هذه هي الطريقة التي اكتشف بها سايمون–ديني بواسون Simeon-Denis Poisson– التوزيع الخاص به.

التوزيع الاحتمالي المتصل Continuous probability distributions

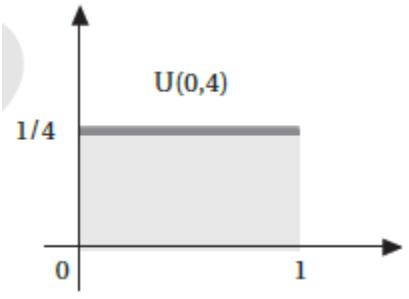
إذا اخترنا شخصا في مدينة ما اختيارا عشوائيا وقمنا بقياس طوله لن نتمكن في هذه الحالة من تطبيق التوزيعات السابقة مثل توزيع ذات الحدين أو توزيع بواسون. المشكلة هنا تكمن في أن الأطوال كميات متصلة وليست منفصلة؛أي لا تشكل أرقاما صحيحة منفصلة: 1،2،3. إنما نستخدم مدى من الأرقام المتصلة. يمكن تطبيق التوزيع المتصل في مثل تلك الحالات(البيانات عبارة عن أطوال أشخاص مثلا). ويطبق عن طريق رسم منحنى لكل حالة، ويسمى هذا المنحنى دالة كثافة الاحتمال.

في التوزيع المنفصل نحسب احتمال وقوع X في مدى معين عن طريق القيام بعملية الجمع لكل الاحتمالات المناظرة، أما في التوزيع المتصل نقوم بعملية التكامل على المنحنى على امتداد المدى المطلوب، وبالتالي يكون احتمال أن X يقع بين العددين 4، 6 يمكن تمثيله بالمساحة تحت المنحنى في هذا المدى.

في التوزيع المنفصل لابد أن يكون المجموع عند جمع كل الاحتمالات يساوي الواحد الصحيح. أما في التوزيع المتصل لابد أن تكون المساحة تحت المنحنى تساوي الواحد الصحيح أبسط توزيع متصل هو الواحد المنتظم (uniform one).

التوزيع الذي يشغل الأهمية الكبرى وبمثابة القلب النابض للنظرية الحديثة للاحتمال هو التوزيع الطبيعي.

التوزيع المنتظم Uniform distribution



تخيل لعبة نحلة دوارة محيطها طوله 4سم وعليها علامات مرقمة من 1 إلى 4. قمنا بتدويرها مرة ثم سجلنا النقطة الموجودة على محيط النحلة والتي تكون ملامسة للأرض عند استقرار النحلة تماما. رقم النقطة سيكون واقعا بين 1، و4 (لا يشترط أن يكون رقما صحيحا).

وبفرض أن النحلة منتظمة، إذا يمكن التعبير عن هذه المسألة باستخدام متغير عشوائي X له توزيع منتظم.

يتحدد التوزيع المنتظم بنقطتيه المتطرفتين (في هذه الحالة (0،4))، وجميع الفترات الواقعة بين هاتين النقطتين ولها نفس الطول لها احتمالات متساوية. بالتالي يكون احتمال سقوط الاسطوانة بين 1.1 و 1.3 مساويا لاحتمال سقوطها بين 3.7، 3.9 على سبيل المثال.

وبشكل عام: إذا كان $X \sim U(a, b)$ تكون دالة كثافة الاحتمال لها قيمة ثابتة بين a وb تعطى بالعلاقة $\frac{1}{a-b}$ وقيمتها صفر فيما عدا ذلك. في المثال السابق، ينتج عن تطبيق ذلك الحصول على رسم بياني ارتفاعه $\frac{1}{4}$. المثال النمطي لهذا النوع من التوزيع هو التوزيع المنتظم القياسي $U(o, 1)$. توقع التوزيع المنتظم العام يحسب من العلاقة $\frac{a+b}{2}$ وتباينه يحسب من العلاقة $\frac{1}{12}(b-a)^2$.

التوزيع الطبيعي Normal distribution

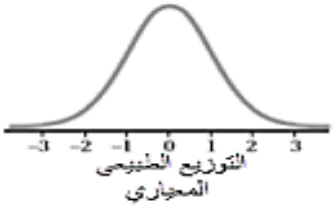
بعيدا عن علم الرياضيات، يعرف منحنى التوزيع الطبيعي باسم منحنى الجرس- bell curve.

عرف لأول مرة في العمل الذي قدمه إبراهام ديموافر (Abraham De Moivre) عام 1756م تحت اسم مذهب الفرص(The Doctrine of Chance) على خلفية عمل ليونارد أويلر(Leonhard Euler's) على الدالة الأسية، لتحديد هذا النوع من التوزيع نحتاج معرفة كميتين: متوسط التوزيع μ، الذي يعين مركز التوزيع بدقة، وتباينه σ^2 الذي يحدد تشتت التوزيع. وهناك مجموعة كبيرة من الظواهر يمكن وضع نموذج لها باستخدام التوزيع الطبيعي للمتغير العشوائي $X \sim N(\mu, \sigma^2)$ X X عن طريق اختيار قيم مناسبة لـ μ وσ^2، من هذه الظواهر: طول مجموعة من النساء البالغات، ونتائج اختبار، وسرعة دوران الكواكب.

الشكل الأساسي لمنحنى الجرس يعطى من بالمعادلة $y = e^{-x^2}$ ولكن لابد من إعادة تحجيمه حتى تصبح المساحة تحت المنحنى مساوية 1 عن طريق نقل μ إلى المركز، وتوسيع

امتداد المنحنى تبعا لقيمة a^{-2} و بدمج ما سبق معا تصبح دالة كثافة الاحتمال:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



في حالة التوزيع الطبيعي المعياري N(0,1) يمكن تبسيط ذلك قليلا إلى $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

وهذه النسخة المعيارية هي ما نحتاجه حقا لأن أي توزيع طبيعي يمكن جعله توزيعا طبيعيا معياريا.

إذا كان $X \sim N(\mu, \sigma^2)$ فإنه يمكن جعله معياريًا عن طريق تعريف $\gamma = \frac{X-\mu}{\sigma}$، ومن ثم $Y \sim N(0,1)$ التوزيع المعياري هو بمثابة التوزيع الأم لكل توزيعات الاحتمال، بمعناه الدقيق المعطى من مبرهنة الحد المركزية the central limit theorem.

المتغيرات العشوائية المتطابقة المستقلة

Independent, identical random variables

تتضمن العديد من الحالات في نظرية الاحتمال متوالية من المتغيرات العشوائية المتطابقة المستقلة ... X_1, X_2, X_3, X_4 . المثال القياسي على ذلك هو القيام بتجربة واحدة عدة مرات بحيث تكون كل مرة مطابقة لما قبلها ، لكن نواتج التجربة في كل مرة تكون مستقلة عن المرات الأخرى مثال: الرمي المتكرر لحجر نرد. (اختيار ورقة لعب من مجموعة أوراق فقط إذا تم تبديل الورقة وإعادة خلط الأوراق كل تجربة).

والشائع هو أن نحسب متوسط أول عدد n مرات من تكرار التجارب. نقوم بذلك عن طريق تعريف متغير عشوائي جديد.

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

حيث Yn تمثل متوسط عدد نواتج التجارب n الأولى (متوسط نتائج أول n رمية لحجر النرد) ويصف قانون الأعداد الكبيرة the law of large numbers ، ومبرهنة الحد المركزية the central limit theorem هذا المتغير.

قانون الأعداد الكبيرة Law of large numbers

ارم حجر نرد منتظم 10 مرات، أو 100 مرة، أو 1000 مرة، واحسب متوسط الناتج في كل مرة. ما الذي تتوقع حدوثه؟ يتنبأ قانون الأعداد الكبيرة أنه كلما كبرت العينة، يجب أن نتوقع أن متوسط نتائج هذه العينة سوف يقترب أكثر من المتوسط النظري الذي قيمته 3.5.

كان جيكوب بيرنولي–Jacob Bernoulli هو أول من يضع ذلك على هيئة نظرية محكمة عام 1713م على الرغم من ظهور أشكال غير رسمية لهذه النظرية قبل عدة سنوات من ذلك. تشير النظرية إلى متوالية من المتغيرات العشوائية المتطابقة المستقلة لكل منها متوسط μ، ومن ثم نضع تعريفا لمتغيرات عشوائية جديدة

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

يجزم قانون الأعداد الكبيرة بأن كلما ازدادت n،اقتربت قيمة المتغير العشوائي Yn للرقم المحدد μ.

صقلت مبرهنة الحد المركزية هذا القانون ولكن القانون يحظي بتطبيقية أوسع؛ حيث تتطلب مبرهنة الحد المركزية وضع افتراض بشأن تباين المتغير Xi.

مبرهنة الحد المركزية Central limit theorem

عام 1733م استخدم إبراهام ديموافر Abraham De Moivre التوزيع الطبيعي لوضع نموذج للعدد الكلي لظهور الصورة في تجربة رمي قطعة نقود عدد كبير من المرات. يبدو أن هناك خطأ ما: عدد رميات قطعة النقود منفصل وليس مستمر فيكون نموذج توزيع ذات الحدين هو الملائم هنا وليس التوزيع الطبيعي!

ومع ذلك لم يرتكب ديموافر خطأ ولو بسيط، على العكس كان ما فعله هو أول شرارة لظهور نتيجة أساسية في نظرية الاحتمال هي: نظرية الحد المركزية.

بفرض وضع نموذج لتجربة ما باستخدام متغير عشوائي X وكان المتوسط μ، والتباين a^2، لا يهم معرفة توزيع X ، ولكن الأمر المهم هو معرفة المعلومات السابقة. قد

يكون توزيعا منتظما، أو توزيع بواسون، أو حتى أي توزيع لم يتم اكتشاف نوعه بعد. تقول نظرية الحد المركزية أنه إذا كررت التجربة عدة مرات ستكون النتيجة المتوسطة معطاه باستخدام توزيع طبيعي، ولمزيد من الدقة تتوافق مرات تكرار التجربة مع متتالية من المتغيرات العشوائية المستقلة المتطابقة $X_1\ X_2\ X_3\ X_4 \ldots$ ومن ثم يمكننا أخذ متوسطات العينة

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

تخبرنا مبرهنة الحد المركزية أنه كلما أصبحت n قيمة كبيرة جدا أصبح التوزيع Yn تقريبا $N(\mu, \frac{\sigma^2}{n})$ وتصبح الأمور أكثر دقة عند جعل Yn قياسيا بتعريف $Z_n = \frac{y_n - \mu}{a/\sqrt{n}}$.

بذلك نقول أن كلما زادتn تصبح المتغيرات العشوائية Zn أقرب إلى التوزيع الطبيعي المعياري $N(0,1)$.

مغالطة المقامر Gambler's Fallacy

قانون المتوسطات هو مصطلح قد يكون معروفا خارج علم الرياضيات، ولكنك لن تجد أثرا لهذا المصطلح في أي كتاب يتناول نظرية الاحتمال. عندما يكون استعمال هذا القانون صالحا يشير إلى قانون الأرقام الكبيرة، أما إذا كان خاطئا يصبح مثالا على مغالطة المقامر. بفرض أن مقامر رأى أن اللون الأسود كان اللون الرابح في المقامرة 6 مرات متتابعة فقد يعتقد أن بعد المرة السادسة سيحين دور الأحمر ليكون هو اللون الرابح وبالتالي يكون احتمال ظهوره في لعبة الروليت أكبر في المرات التالية. إذا كانت هناك محاولات متتالية (دوران قطعة النقود أو عجلة المقامرة) وعرف عنها أنها مستقلة، فستصبح مغالطة المقامر – بالتعريف خاطئة.

مغالطة المقامر تظهر كنتيجة لوجود خطأ في نظرية الأعداد الكبيرة. الخطأ هو أن هذا القانون يتنبأ بتنبؤات احتمالية عن سلوك متوسط على مدى بعيد، ولا تقدم تنبؤات عن نتائج تجارب فردية بالطبع وجود عدد كبير من النتائج مستقلة لازال يستحق أن يؤخذ في عين الاعتبار حيث أنه دليل على أن التجارب ليست كما تبدو (إما أن الاحتمالات لاتتحقق أو الأحداث ليست مستقلة تماما).

العمليات العشوائية STOCHASTIC PROCESSES

العمليات العشوائية Stochastic Processes

ابدأ المشي من منتصف طريق ما، ثم قم بإلقاء قطعة نقود منتظمة. إذا كانت نتيجة الرمية ظهور صورة امش مسافة متر في اتجاه الشمال، أما إذا كانت كتابة امش مسافة متر في اتجاه الجنوب. ثم ألقها مرة أخرى. إلى أين سينتهي بك المسير بعد إلقائها 10 مرات أو 100 مرة؟ هذه التجربة مثال على سير عشوائي.

بالنسبة للسير العشوائي المعتمد على اتجاهين، لنأخذ على سبيل المثال ولاية مانهاتن المعروفة بشوارعها التي على شكل شبكة مربعة. عند كل مفترق طريق يكون لديك احتمال متساو للسير شمالا أو جنوبا أو شرقا أو غربا مربع واحد (وضعنا نموذج للمدينة بحيث نفرض أنها شبكة لا نهائية مع إهمال إمكانية الوصول لحوافها)، وبالمثل يمكن تعريف سير عشوائي في مستوى ثلاثي الأبعاد أو أي مستوى لانهائي (لأن المسألة تكون تافهة إذا كانت المستويات نهائية).

السير العشوائي هو أبسط مثال على العمليات العشوائية، وهي العمليات التي تتغير بتغير الزمن استنادا إلى قواعد احتمالية وليس علاقات محددة. تعتبر سلاسل ماركوف– Markov chains والحركة البراونية –Brownian motion أمثلة أكثر تعقيدا.

السير العشوائي لبوليا Polya's random walks

صاغ جورج بوليا George Polya مصطلح السير العشوائي عام 1921م بعد دراسته لحالتي السير العشوائي أحادي البعد، وثنائي الأبعاد الوارد ذكرهما في المثالين السابقين. كان سؤاله كالتالي: في البداية اختر نقطة على الرسم البياني. الآن احسب احتمالية أن يصل إليها في النهاية الشخص السائر سيرا عشوائيا؟ السؤال الأسهل الذي يفضي إلى نفس الأمر: ما هو احتمال أن يعود السائر في نهاية المطاف إلى نقطة البداية؟ أوضح بوليا أن النتيجة هي 1 في كلا الحالتين وجعل ذلك تأكيدا افتراضيا.

تعرف حالة السير أحادي البعد باسم دمار المقامر gambler's ruin. المقامر الذي

يلعب لعبة منتظمة عشوائية ضد نادي القمار لديه احتمال 1 لخسارة كل عملاته في نهاية اللعبة. قد يبدو ذلك غير مثير للدهشة، لكن بوليا وضح أن بزيادة عدد الأبعاد يصبح الأمر غير متحقق، فالسير العشوائي على شبكة ثلاثية الأبعاد تكون احتمالية الرجوع فيه إلى نقطة البداية أقل، وبعد ذلك تثبت قيمة هذا الاحتمال عند حوالي 0.34، فالسير العشوائي ذو الأبعاد الأكثر لا يغطي الشبكة كلها بل يظهر سلوكا كسريا مثيرا للدهشة كلما زادت الأبعاد.

سلاسل ماركوف Markov chains

في كل مرحلة من مراحل السير العشوائي لمانهاتن تقوم بإلقاء قطعة نقود لمعرفة الاتجاه التالي الذي عليك أخذه. تعتبر نظرية الاحتمال أن رمي قطعة النقود يمثل بمتغير عشوائي من نوع بسيط. سلسة ماركوف هي متتالية من المتغيرات العشوائية مثل السير العشوائي، لكن يكمن الفرق في أن هذه المتغيرات العشوائية قد تكون من نوع أكثر تعقيدا مثل السير العشوائي على شبكة بها ناقلات أو أي أفخاخ خداعية أخرى. عبر عالم الاحتمالات أندري ماركوف في القرن التاسع عشر عن السمات التي تحدد سلسلة ماركوف. سمات هذه السلسلة أن التوزيع الاحتمالي في كل مرحلة يعتمد فقط على الحاضر وليس الماضي (في السير العشوائي المهم النقطة التي وصلت إليها، أما كيفية وصولك إلى تلك النقطة فهو مهمل).

تمثل سلاسل ماركوف إطارا ممتازا لوضع نماذج للعديد من الظواهر بما فيها: التغيرات السكانية، وتقلبات البورصة. تحديد السلوك النهائي لعملية ماركوف يمثل مشكلة عميقة كما توضح شروحات بوليا للسير العشوائي ثلاثي الأبعاد.

النظرية الحركية للغازات Kinetic theory of heat

كان روبرت براون (Robert Brown) عالم نبات، ورائدا في الفحص الميكروسكوبي في مجال العلوم الحيوية. في عام 1827م وجه مجهره نحو حبوب لقاح زهرة الربيع العالقة في الماء. كانت الأجزاء الصغيرة من المادة تطفو على الماء وتندفع اندفاعا عشوائيا، عرف ذلك فيما بعد باسم الحركة البراونية.

اعتقد براون في بداية الأمر أن هذه الجسيمات عبارة عن مخلوقات حية ضئيلة، ولكن

المزيد من التقصي كشف عن أن نفس الحركة المميزة غير المنتظمة ظهرت في مسحوق صخري مذاب جيدا في الماء.

عام 1905م أدرك العالم آينشتاين- أن هذه الجسيمات تدفع بواسطة جزيئات الماء، وهي ضئيلة جدا لايمكن رؤيتها. ومن الجدير بالذكر أن بزيادة درجة حرارة الماء تتحرك الجزيئات المرئية بشكل أسرع. أدرك آينشتاين أن هذا دليل قوي غير مباشر على النظرية الجزيئية للحرارة- وكما نعلم ، الطاقة الحرارية في المادة ليست إلا طاقة حركية للجزيئات المكونة لها.

الحركة البراونية Brownian motion

كان لابد من وجود نموذج رياضي للحركة البراونية للجزيئات حتى يتسنى لنا تجسيد تفاصيل عمل آينشتاين على النظرية الحركية للغازات، وحيث أن تغيير مسار الجزيء تغيير عشوائي ومستقل عن حركته السابقة فهو يتشابه مع العملية العشوائية كما في سلسلة ماركوف. ولكن في السير العشوائي وسلاسل ماركوف يكون الوقت على هيئة خطوات منفصلة؛ حيث تحدث كل خطوة في العملية بعد مدة ثابتة من الزمن، أما في الحركة البراونية يغير الجسيم من اتجاهه باستمرار فيبدو المسار على هيئة سير عشوائي مصغر بحيث تؤول خطوات الانتقال إلى صفر.

يمثل السير العشوائي بمتوالية من المتغيرات العشوائية ... $X_1, X_2, X_3, X_4,$؛ أي مجموعة من Xi حيث i رقم طبيعي. وعلى النقيض من ذلك ، تمثل حركة براون بمجموعة متصلة من المتغيرات العشوائية ؛ أي مجموعة من Xi حيث i رقم حقيقي.

ماذا يحدث عند توسيع هذا النظام؟ بين آينشتين أن بعد أي مدة زمنية يمثل موضع الجسيم بتوزيع طبيعي ثلاثي الأبعاد (موضعه في كل بعد موزع توزيعا طبيعيا، والأبعاد الثلاثة مستقلة).

التشفير CRYPTOGRAPHY

التشفير أحادي الحرف Monoalphabetic encryption

من أبسط الطرق المستخدمة لتشفير رسالة هو إعادة ترتيب الحروف قبل كتابة الرسالة. على سبيل المثال: نظام تشفير قائم على ترتيب الحروف على لوحة المفاتيح:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	Y	Z
Q	W	E	R	T	Y	U	I	O	P	A	S	D	F	G	H	J	K	L	Z	X	C	V	B	N	M

يمكننا استخدام هذا المفتاح في تشفير رسالتنا، ولكن قبل البدء في التشفير : تعرف الرسالة باسم (النص الصريح) وستكتب بحروف صغيرة. لنفرض أن رسالتك تقول: "قابلني في الحديقة عند الثالثة صباحا–'meet me in the park at three a.m. ينتج النص المشفر (الذي سيكتب بحروف كبيرة) عن طريق تبديل الحروف طبقا للجدول بالأعلى 'DTTZ DT OF ZIT HQKA QZ ZIKTT Q.D.'

وبمجرد إرسال ذلك إلى جهة الاتصال المطلوبة، تقوم بدورها بقراءة الرسالة، باستخدام نفس المفاتيح. وبالتأكيد ليس هناك أي سبب وراء اختيار حروف الهجاء فقط؛ أي أن أي 26 رمز ستكون جيدة. ولجعل الجملة أقل إرشادا قد نود حذف بعض علامات الترقيم والمسافات.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
√	=	∩	÷	א	∃	Σ	≠	→	∫	∈	∞	x	∪	±	e	$\frac{dy}{dx}$	Π	1	π	i	∀	θ	⊆	∅	×

$$' \rightarrow \neq \sqrt{\forall א \pi} \neq אx\sqrt{\pi א} II \rightarrow \sqrt{\infty} 1 \emptyset \pm i\ II א \frac{dy}{dx} i \rightarrow II א'$$

تحليل الشفرات Cryptanalysis

لنفرض أنك قد وقع في يديك رسالة مشفرة مرسلة لعدو لك:

WKRKRPUEBRXEUGRJURJFBGDRFRBKGFRGBGURPBJRXFOKGRPURZ
UGRXAKIJRPK

AOFXOFVUDIJUXKIFJUGKRAKQQKZULXKIJEKGRFERBDFQUSFDPUZB
QQPFTUFZPB

RUEFJGFRBKGBGPUJPFGLXKIJEKLUZKJLBDDBDXYPIDDPUZBQQWBT
UXKIFVUXR
PUOFRUJBFQDFJUBGMKSGIOMUJDBSUBWPRABTULUDRJKXRPBDOU
DDFWUFAR
UJXKIPFTUOUOKJBDULBRLKGKREKGRFEROUFWFBGUGLKAOUDDFWU

كيف يمكن المحاولة في فك شفرتها؟ هذا هو السؤال الذي يسأله علم التشفير.

لنفرض أن المرسل استخدم أسلوب أحادي الحرف (الطريقة الأم التي اكتشفت في القرن 19 على يد أبوالكندي (Abu al-Kindi). أساس التحليل التكراري هو ملاحظة الحروف التي ليس لها نفس الشيوع. الخطوة الأولى تحليل أكثر الحروف شيوعا في النص المشفر

U	R	F	D	B	G	D	J	P
36	29	26	25	23	19	18	17	15

الفكرة الأساسية: هي محاولة إبدال الأرقام بالحروف الأكثر تكرارا التي تنشأ في الانجليزي. وهي على الترتيب مجموعة في ETAOINSHRDLU.

تحليل التكرار Frequency analysis

في المثال السابق يمكن أن نبدأ بتبديل الحرفين الإنجليزيين الأقل شيوعا t،e بأكثر حرفين شيوعا U، R، وهو ما يعطينا:

WKtKtPeEBtXEeGtJetJFBGDtFtBKGFtGBGetPBJtXFOKGtPetZeGtXAKIJt
PKAOFXOFVeDIJeXKIFJeGKtAKQQKZeLXKIJEKGtFEtBDFQeSFDPeZBQQ
PFTeFZPBteEFJGFtBKGBGPeJPFGLXKIJEKLeZKJLBDDBDXYPIDDPeZB
QQWBTeXKIFVeXtPeOFteJBFQDFJeBGMKSGIOMeJDBSeBWPtABTeLeDtJ
KXtPBDOeDDFWeFAteJXKIPFTeOeOKJBDeLBtLKGKtEKGtFEtOeFWFB
GeGLKAOeDDFWe

الآن يمكننا فتح خط هجومي آخر. في بعض المواضع يكون الحرف t في الرسالة الصريحة متبوعا بالحرف المشفر P، ومن معرفتنا باللغة الإنجليزية يبدو أنه P تمثل h، وهو ما يعطينا:

WKtKtheEBtXEeGtJetJFBGDtFtBKGFtGBGethBJtXFOKGthetZeGtXAKIJt
hKAOFXOFVeDUeXKIFJeGKtAKQQKZeLXKIJEKGtFEtBDFQeSFDheZBQ
QhFTeFZhBteEFJGFtBKGBGheJhFGLXKIJEKLeZKJLBDDBDXYhIDDheZ
BQQWBTeXKIFVeXtheOFteJBFQDFJeBGMKSGIOMeJDBSeBWhtABTeLe
DtJKXthBDOeDDFWeFAteJXKIhFTeOeOKJBDeLBtLKGKtEKGtFEtOeFW
FBGeGLKAOeDDFW

ومن خلال دمج تحليلات تكرار حروف مختلفة أو تكرار مجموعة من الحروف معا، وذكاء الحل، واستخدام مبدأ التجربة والخطأ يصبح لدينا قدرة على تحقيق تطور أكبر.

ومن الملائم ذكر تحذير مهم بهذا الصدد: تحليل التكرار ليس علما قائما بذاته، وهو يعمل بشكل أفضل مع النصوص الكبيرة. هذا المثال القصير إلى حد ما مفتعل لكن كل ما يراد به هو شرح الأساليب الأساسية.

إيتون شردلو ETAOIN SHRDLU

يعتمد تحليل الكرار على معرفة التكرارات النسبية بين الحروف المختلفة في اللغة الإنجليزية لاشك في أن هذه مجرد متوسطات ولن يكون عددا دقيقا في أي نص.

الحرف	معدل ظهوره كل 100 حرف	الحرف	معدل ظهوره كل 100 حرف
E	12.7	m	2.4
T	9.1	w	2.4
A	8.2	f	2.2
O	7.5	g	2.0
I	7.0	y	2.0
N	6.7	p	1.9
S	6.3	b	1.5
H	6.1	v	1.0
R	6.0	k	0.8
D	4.3	j	0.2
L	4.0	x	0.2
U	2.8	q	0.1
C	2.8	z	0.1

تجمع العبارة ETAOIN SHRDLU الاثني عشر حرفا الأولى مرتبة حسب تكراريتها، وقد كان ذلك معروفا في العصر الذي ظهرت فيه آلات طباعة اللينوتيب[1] حيث رتبت الحروف على لوحة المفاتيح حسب تكراريتها تقريبا. وأحيانا كانت تظهر هذه العبارة في الجرائد عن طريق الخطأ.

تحليل التكرار ليس مفيدا للحروف المنفردة فحسب، بل مفيد أيضاً مع مجموعات الحروف التي تظهر معا (مثل th التي هي أكثر شيوعا من qz)

حرفان يمثلان صوتا منفردا	معدل مرات الظهور كل 2000 حرف	حرفان يمثلان صوتا منفردا	معدل مرات الظهور كل 2000 حرف	حرفان يمثلان صوتا منفردا	معدل مرات الظهور كل 2000 حرف
Th	50	at	25	st	20
Er	40	en	25	io^^B	18
On	39	es	25	le	18
An	38	of	25	^ is	17
Re	36	or	25	ou	17
He	33	nt	24	ar	16
In	31	ea	22	as	16
Ed	30	ti	22	de	16
Nd	30	to	22	rt	16
Ha	26	it	20	ve	16

التشفير المتعدد للأحرف Polyalphabetic encryption، الأكواد Codes، وأخطاء التهجئة Spelling mistakes

هناك أساليب عديدة تزيد من صعوبة فك التشفير أحادي الحرف. أحد هذه الأساليب هو استخدام التشفير متعدد الأحرف وفيه يكون هناك أكثر من طريقة لتشفير الحرف الواحد، وبذلك قد يستخدم المفتاح الواحد 52 حرفا هجائيا وكل حرف من الشفرة مشفر

(1) قطعة معدنية تحتوى على قوالب معدنية تمثل كل الحروف المستعملة منضدة بجوار بعضها بعضا، وقد أطلق عليها اسم "خط الحروف الطباعية-Linotype .

باختيار رمزين. لمزيد من التعقيد، نستخدم الرموز الوهمية–dummy symbols وهي رموز ليس لها معنى وسيقوم المتلقي (الموجه إليه الرسالة المشفرة) بحذفها ولكنها ستشوش أي متلقي دخيل:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	dummy
S	L	6	M	D	R	{	E	Q	W	\	A	@	B	K	7	J	3	T	C	O	?	G	4	P	H	X
Z	!	(	5	%	*	1	)	–	9	+	F	£	$	2	N	Y	~	I	;	8	^	U	#	,	V	}

وإحدى الوسائل القديمة هي إضافة مستويات عليا من التشفير. حتى يومنا هذا نقوم بتشفير الرسائل باستخدام رموز مختلفة ترمز إلى حروف الرسالة، وهذا ما يسمى الشفرة، أما الكود– فهو استخدام كلمات بشكل مقنع، على سبيل المثال:

بنك	دولار	سيارة	رجل شرطة
نعامة	رخام	طبلة	صلصة اللحم

أما الوسيلة الثالثة فهي استخدام أخطاء تهجئة متعمدة في الرسالة، وذلك لمنع التحليل التكراري من اكتشاف الأمر. الثلاث وسائل السابقة تعمل على جعل فك تشفير الرسائل أصعب منه في التشفير أحادي الحرف البسيط:

{DXC}C;EX28HXS$}B5@ZX3~!ATRX~K£X}@KXI;3,(E{}P?CX2K1~S
A,P$5X38}£2O;V}M%’

لوحة المرة الواحدة One - time pad

نقطة ضعف التشفير أحادي الحرف هو أن الحرف يتم تشفيره بنفس الطريقة كل مرة مما يفتح الباب على مصراعيه أمام التحليل التكراري وفك التشفير. خففت حدة هذه المشكلة باستخدام وسائل أخرى مثل التشفير المتعدد للأحرف، والتشفير باستخدام الأخطاء الهجائية، وغيرهما من الوسائل. ولكن في نهاية المطاف قد يتمكن أي محلل شفرات خبير من اجتياز تلك العقبات وخاصة إذا كان مزودا بجهاز كمبيوتر لكشف الإمكانات المختلفة للحروف.

لوحة المرة الواحدة هي وسيلة بديلة تعمل على استخدام سلسلة من الحروف كمفتاح لها. لنفرض أن المفتاح يبدأ بكلمة (mathematical رياضي)

أولا: تحول حروف النص غير المشفر إلى أرقام تبعا لترتيبهم الهجائي.

النص غير المشفر(plain text)

a	b	o	r	t	m	i	s	s	i	o	n
1	2	15	18	20	13	9	19	19	9	15	14

المفتاح (Key)

a	m	t	h	e	m	a	t	i	c	a	L
13	1	20	8	5	13	1	20	9	3	1	12

ونحصل على النص المشفر(Ciphertext) بجمع كل رقمين متناظرين ثم إعادة تحويلها إلى حروف، وإذا كان الناتج أكبر من 26 نطرح من الناتج 26. بمعنى أن عملية الجمع تتم بتردد 26 (Modulu 26) (انظر العمليات الحسابية النمطية).

النص المشفر (Ciphertext)

14	3	9	26	25	26	10	13	2	12	16	26
N	C	I	Z	Y	Z	J	M	B	L	P	Z

يقوم المتلقي بعد ذلك بفك تشفير الرسالة عن طريق القيام بعملية عكسية لما تم شرحه في الأعلى (طالما يعرف مفتاح الشفرة).

هذا النوع من التشفير يكافئ تشفير حروف متتالية تبعا لشفرات أحادية مختلفة. يرجع المستوى الأمني العالي لطريقة تشفير لوحة المرة الواحدة وتسميتها بهذا الاسم إلى أن كل مفتاح يستخدم لتشفير رسالة واحدة فقط ثم يتم التخلص منه لذلك قد تكون هناك لوحات متطابقة وفيها مفتاح في كل صفحة ونقلب صفحة جديدة مع كل رسالة.

مبدئيا لا يمكن اختراق طريقة لوحة المرة الواحدة أبدا، كما أثبت كلود شانون (Claude Shannon) عام 1949م.

هناك العديد من الرسائل غير المشفرة الممكنة والتي لها نفس الطول وبدون المفتاح لن يتمكن أي محلل شفرة عدو من التفرقة بين تلك الرسائل.

المفاتيح العامة Public Keys

نظريا لا يمكن خرق طريقة اللوحة الواحدة إلا أنها تحتاج إلى مفاتيح كثيرة جدا؛ حيث يحتاج المرسل والمتلقي مفتاح جديد لكل رسالة. القيام باستبدال هذه المفاتيح محفوف بالمخاطر الحتمية. في طريقة لوحة المرة الواحدة أو أي طريقة تشفير تقليدية تكون عمليتي التشفير وفكه عمليتين متماثلتين. بمعنى أن كل من المرسل والمتلقي يحتاج أن يكون لديه نفس المفتاح.

أما في المفاتيح العامة فلا وجود لهذا التماثل؛ حيث يصبح المفتاح في هذه الحالة جزأين: الجزء الخاص (Private Part) وهو الذي يحتفظ به المالك، ولا يشاركه مع أحد أبدا، والجزء العام(Public Part) الذي يسمح للجميع بالوصول إليه.

يستطيع أي شخص تشفير رسالة باستخدام المفتاح العام ويرسلها إلى المالك الذي يكون هو الشخص الوحيد الذي في إمكانه فك التشفير لأن ذلك يتطلب مفتاح خاص

بالمثل تخيل أن مؤسسة عمل ما توفر عددا مطلقا من الأقفال المتطابقة غير الموصدة ولكل منها مفتاح وحيد تملكه المؤسسة وحدها. إذا أراد شخص ما أن يرسل لها شيئا يمكنه وضعه في صندوق وقفله بأحد الأقفال. يمكن للمؤسسة فقط أن يفتح هذا القفل فيما بعد.

التشفير باستخدام مفتاح عام (Public key cryptography) هو العمود الفقري لأمن الشبكات الحديث (Modern internet security). يتكون المفتاح من رقميين أوليين كبيرين وليكونا p و q ويتم الإبقاء على سريتهما بينما يصرح عن حاصل ضربهما $p \times q$.

يعتمد أمن هذا النظام على الصعوبة المتأصلة في القيام بالعملية العكسية لهذه العملية؛ أي ما يعرف باسم مسألة تحليل الأعداد الصحيحة (Integer factorization problem).

نظرية المعلومات لشانون Shannon's information theory

كان البحث الذي نشره كلود شانون عام 1948م بعنوان (نظرية رياضية في التواصل A Mathematical Theory of Communication) رائعة من روائع عصر ما بعد الحرب وكان له بصمة في ميلاد موضوع نظرية المعلومات كما أنه أصبح ذا مكانه لا تقدر بثمن منذ النصف الثاني من القرن العشرين. في هذا البحث درس شانون عمليات: التشفير (encoding)، والتراسل وفك تشفير المعلومات (transmitting and deciphering information)، ولكن مساهمته لم تكن طريقة جديدة في عالم التشفير

نبدأ بذكر أنه كان رائدا في استخدام النظام الثنائي (Binary) كلغة أساسية للمعلومات. ثم قام لأول مرة بتحليل الأساس النظري لإرسال المعلومات. وبحث في حدوده وحلل المعدل الأقصى لقدرة النظام على إرسال بيانات و تعتمد الإجابة على مصدر المعلومات وخصوصا على كمية ما تسمى مقياس عشوائية النظام (إنتروبي entropy). وهو مقياس يحدد بدقة تقلب البتات (الخانات في النظام الثنائي) المتتابعة في سيل من الأكواد الثنائية عن طريق وضع نموذج له باستخدام عملية ماركوف(Markov process). اكتشف فيما بعد أن مفهوم الإنتروبي عند شانون يكافئ تعقيد كولموجروف (Kolmogorov complexity).

الحالتان اللاتي وضعهما شانون محل الدراسة هما: الأنظمة الخالية من التشويش (noiseless systems)، والأنظمة المشوشة(noisy systems) (التي يسهل حدوث خطأ فيها) ؛ في الأخيرة، درس الحدود النظرية لقوى تصحيح أخطاء الشفرات (the theoretical limits of the powers of error correcting codes.).

تعقيد كولموجروف Kolmogorov complexity

يتم تمثيل المعلومات في التكنولوجيا الحديثة على شكل تتابعات ثنائية. (آسكي ASCII هو شفرة لتحويل الحروف ورموز أخرى عديدة إلى النظام الثنائي على سبيل المثال). وبعض هذه التتابعات يكون أكثر تعقيدا من الآخرين. يسهل وصف التتابع 111111 حيث أنه يحتوي على كمية ضئيلة من المعلومات. قد يكون تخزين سلسلة مكونة من مليون 1 إهدارا لمساحة القرص، ولذلك تستخدم برامج الأرشفة التي يمكنها ضغط هذا التتابع بشكل كبير عن طريق إعادة تخزين المعلومات على هيئة أوامر لكتابة مليون 1 (111....) بهدف توفير المساحة.

في الستينيات استخدم كل من راس سولومونوف (Ray Solomonoff) وآندري كولموجروف (Andrey Kolmogorov) هذه الفكرة كطريقة لتحديد كمية المعلومات التي تحتويها سلسلة من البتات. تعقيد كوولموجروف لسلسلة ما هو: أقل طول للسلسة يمكن الوصول إليه عند ضغطها. السلاسل التي تحمل معلومات كثيرة تكون غير قابلة للضغط وبذلك يكون تعقيدها عالي. أما السلاسل التي تحمل معلومات قليلة مثل المليون 1 التي ذكرناها يمكن ضغطها بكفاءة وبذلك نقول أن تعقيدها منخفض. تعقيد كولموجروف يكافئ بالأساس مفهوم الإنتروبي لسلسلة من البتات عند شانون.

أكواد تصحيح الأخطاء Error-correcting codes

عند إرسال معلومات خلال قناة مشوشة (Noisy channel) يمكن أن تتسلسل بعض الأخطاء. أكواد تصحيح الأخطاء هي شفرات تمكن الرسالة من تحمل بعض مستويات التلف.

أبسط طريقة لذلك هي التكرار البسيط(Plain repetition): بدلا من إرسال عبارة (COME NOW) نرسل (CCCOOOMMMEEE NNNOOOWWW) إذا أصاب التلف خانة واحدة نستعين بباقي المجموعة: CCCAAANNQNNNOOOTTT3.

أما إذا حدث أكثر من خطأ ستحدث أزمة أخرى. يمكن التغلب على ذلك باستخدام

مجموعات تكرارية أطول وليكن 100 تكرار للحرف لكن سينتج عن ذلك تقليل في سرعة العملية. أوضح شون توضيحا بارزا أن هذه المفاضلة بين الدقة والسرعة ليست حتمية إذ يمكن باستخدام بعض التكنولوجيا الرياضية إيجاد شفرات سريعة وتحقق مستوى الدقة المطلوب. حتى إذا أصاب التلف نسبة 99٪ من الرسالة، قد يظل فك شفرتها بكفاءة شيئا ممكنا. تعتمد إحدى الطرق على المربعات اللاتينية (Latin squares) التي في طبيعتها تعتبر مصححا للأخطاء.

إذا كان أحد مدخلات المربع اللاتيني تالفا يكون من السهل تحديده وتصليحه عن طريق فحص كل الصفوف والأعمدة . وبعض الطرق المتطورة تستخدم البنى الجبرية للمجالات النهائية (the algebraic structures of finite fields).