



THE WORLD BANK



الجلسة السابعة العينات والقوة

Christel Vermeersch

يناير 2008

- الهدف: تقييم البرامج والإجراءات التدخلية
(هل لها أثر أم لا؟)
- نحتاج إلى بيانات كافية لنتمكّن من اكتشاف التغيرات الناجمة عن البرنامج المعني
 - ← الجزء 1: مناقشة القوة
- نحتاج إلى بيانات تعطي صورة صادقة:
 - ← الجزء 2: مناقشة أخذ العينات

الجزء 1: امتلاك بيانات "كافية"

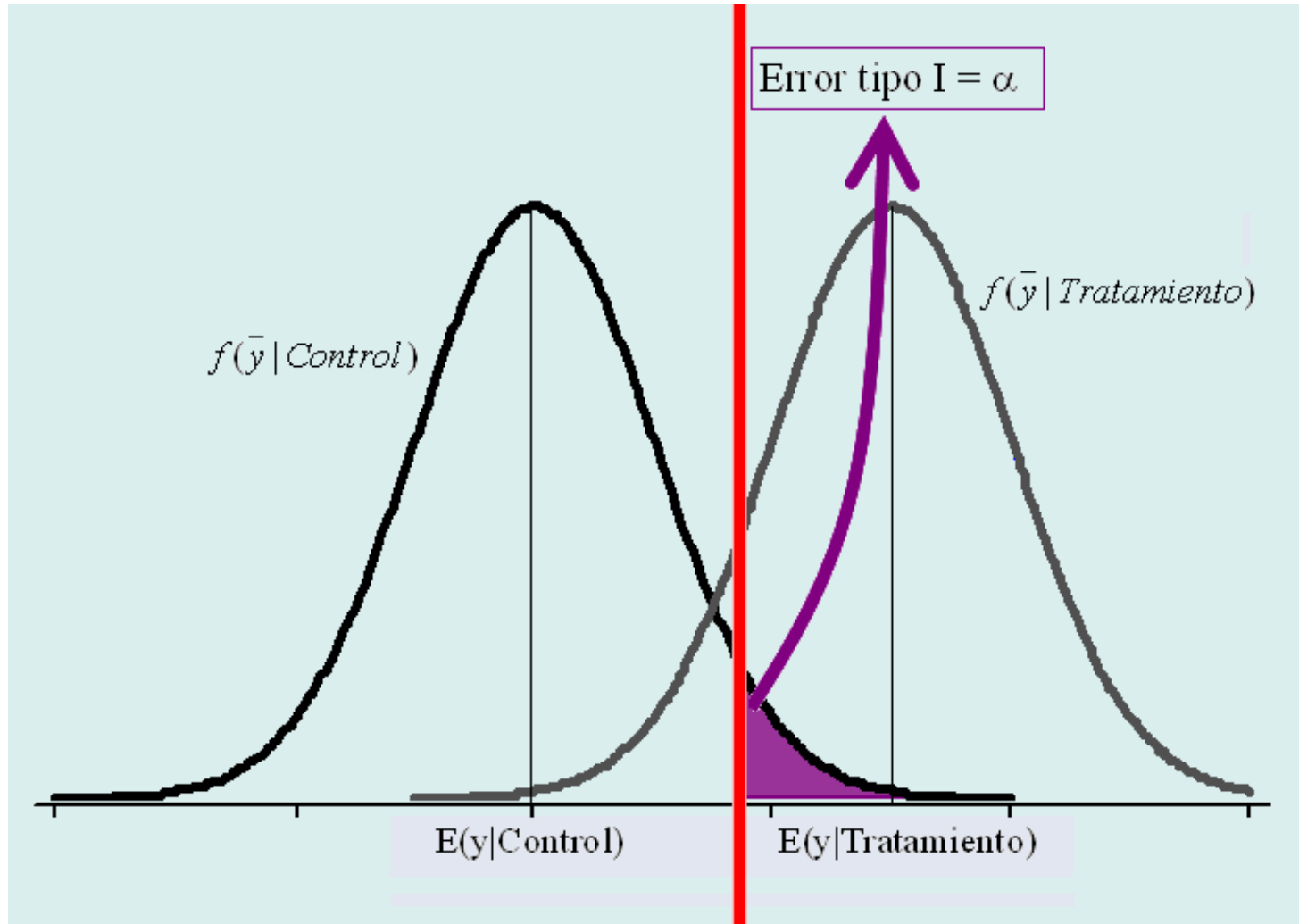
- تقييم البرنامج المعني هو اختبار لأهميته
- مثال: y هو متغير يشير إلى نتائج الاختيار. نسعى لتقدير أثر مضاعفة موازنة المدرسة على نتاج امتحانات الطلبة.

أي ليس للبرنامج أي أثر $H_0: E(y|Treatment) = E(y|Control)$

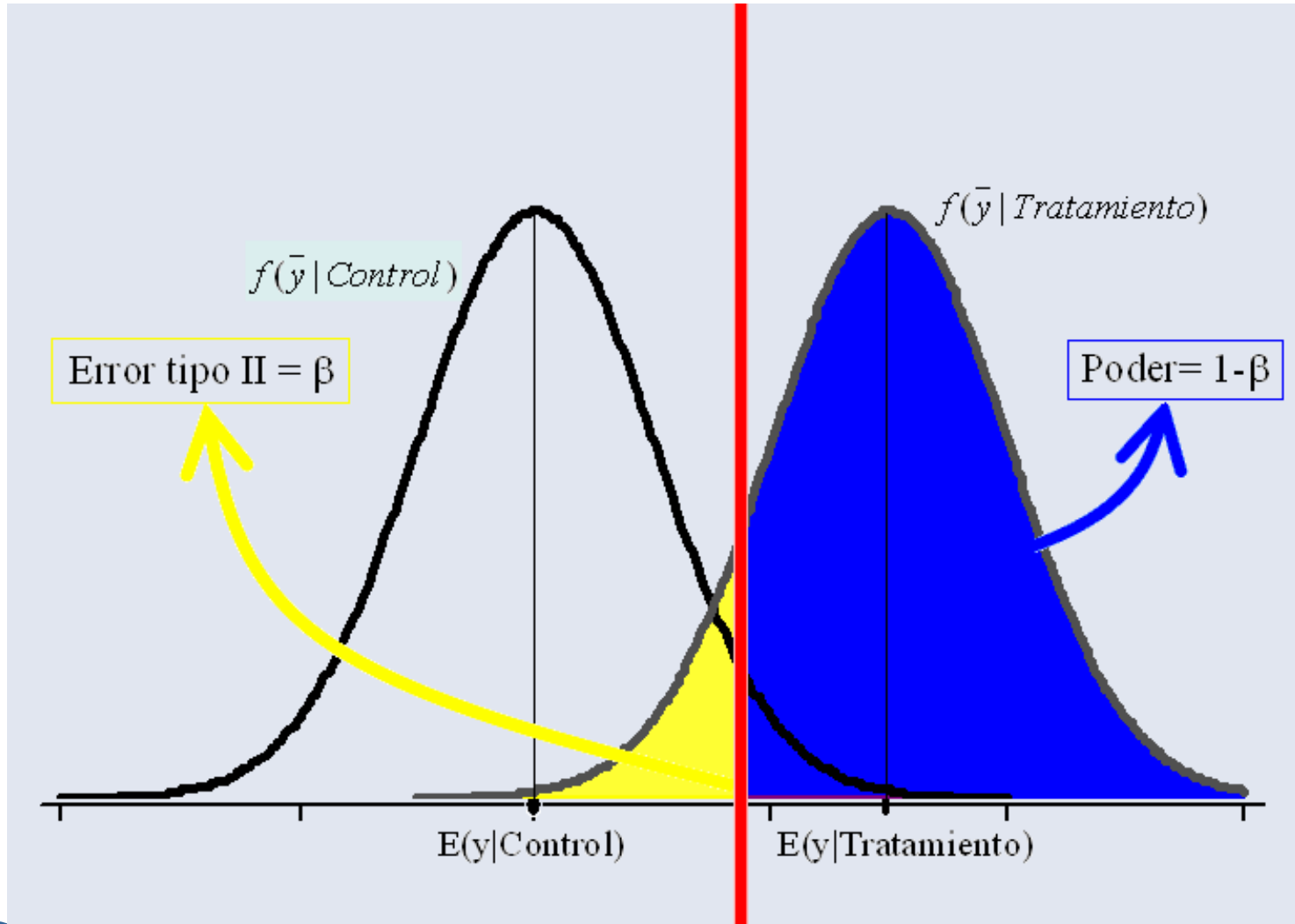
أي للبرنامج أثر إيجابي $H_1: E(y|Treatment) > E(y|Control)$

		للإجراء التدخلي أثر (في الواقع)	
		لا	نعم
اختبار إحصائي	نرفض H_0	خطأ من النوع الأول	لا بأس
	لا نرفض H_0		خطأ من النوع الثاني

الخطأ من النوع الأول



الخطأ من النوع الثاني والقوة



الخطأ من النوع الأول والنوع الثاني

- الخطأ من النوع الأول
 - = رفض فرضية عدم القيمة عندما تكون صحيحة
 - مستوى الأهمية α = احتمال أن نخلص إلى أن للإجراء التدخلي أثر حين لا يكون له أي أثر
 - القيم النموذجية لـ α : 0.01, 0.05, 0.1
- الخطأ من النوع الثاني
 - = عدم رفض عدم القيمة حين يكون غير صحيح
 - القوة = $1 - \beta$
 - = احتمال أن نرفض فرضية عدم القيمة حين تكون غير صحيحة
 - = احتمال أن نرفض فرضية عدم القيمة حين تكون الفرضية البديلة صحيحة
 - = احتمال أن نخلص إلى أن للإجراء التدخلي أثر حين يكون له فعلاً أثر
- الغرض: تعظيم القوة، لمستوى محدد من α

ما الذي يؤثر في قوة الاختبار؟

تزداد قوة الاختبار عندما ...

□ تزداد α (نتوصل إلى تسوية وسط بشأن الخطأ من النوع الأول)

□ عندما يكون الأثر المتوقع أكبر

□ عندما يزداد تغير المؤشر ... كيف يحدث هذا؟؟؟

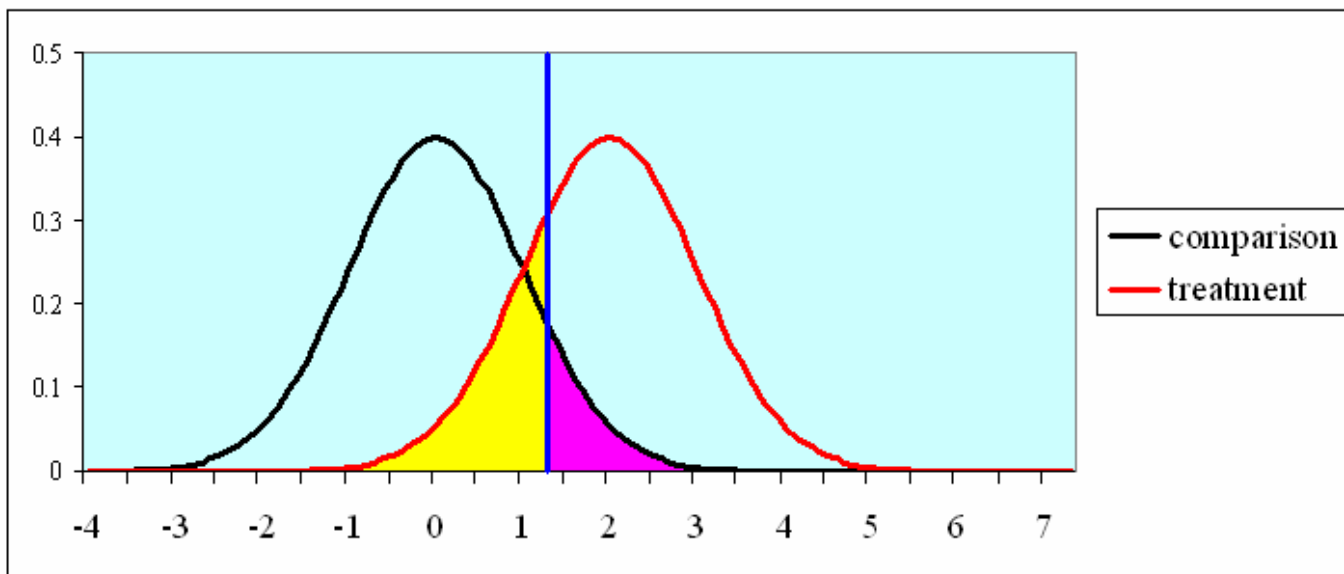
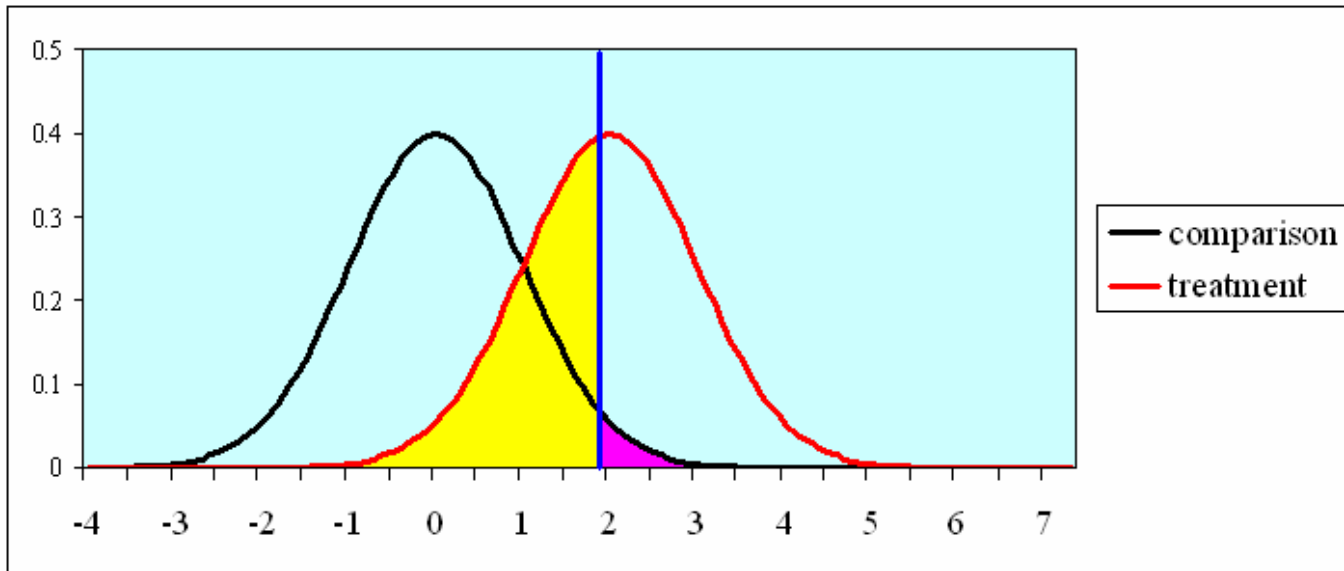
■ ضرورة تذكر أن تغاير المتوسط متناسب مع $1/n$ ، حيث n هي حجم العينة

■ لذلك، من شأن زيادة حجم العينة تخفيض تغاير المتوسط المعني (العديد من المؤشرات هي متوسطات)

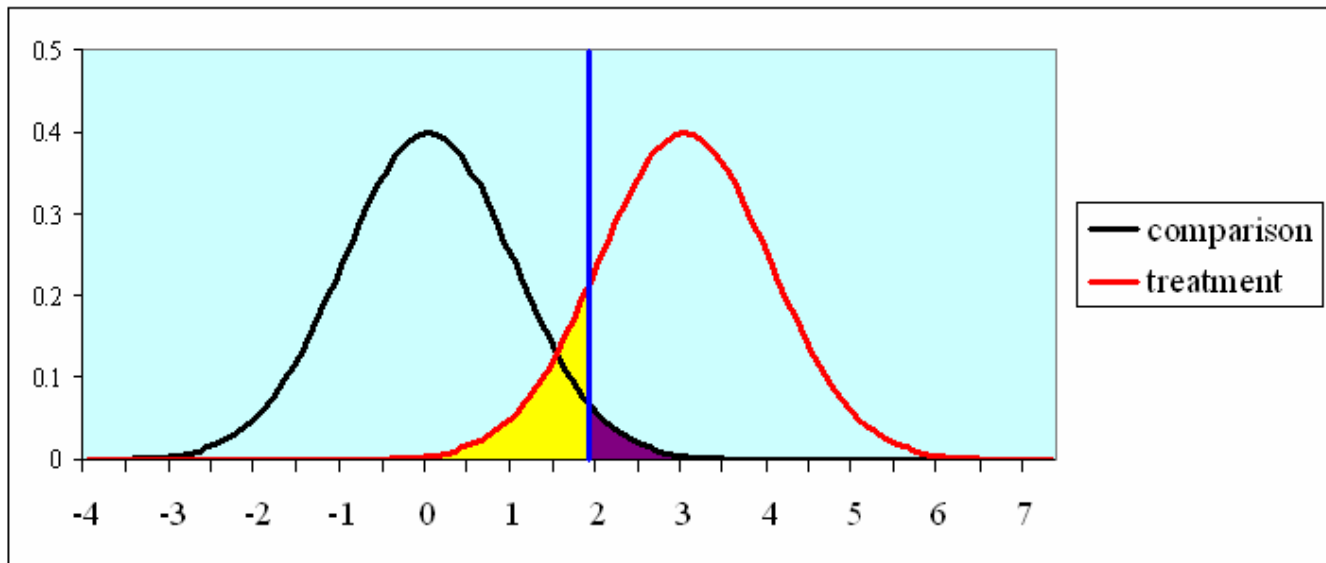
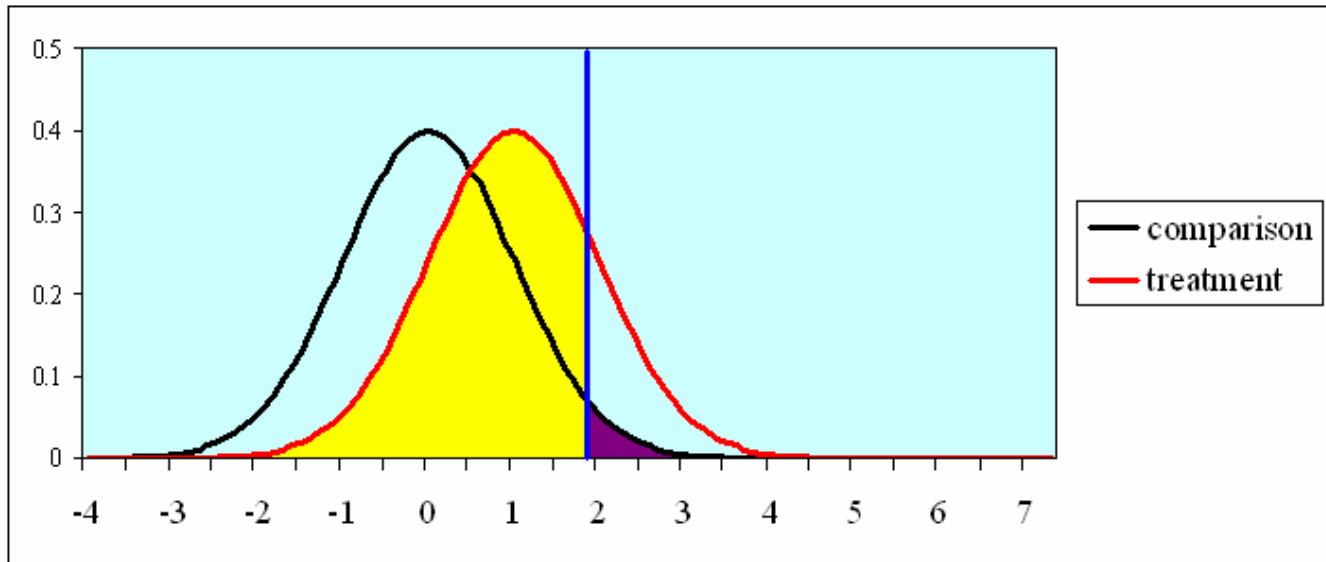
=> لزيادة قوة العينة، حين يكون حجم الأثر ثابتاً والخطأ من النوع الأول ثابت، علينا زيادة حجم العينة.

=> عملياً، نقوم بحساب حجم العينة الضرورية لاكتشاف تغيير في X في المؤشر، مع خطأ من النوع الأول يبلغ 5% وقوة تبلغ 80 أو 90%.

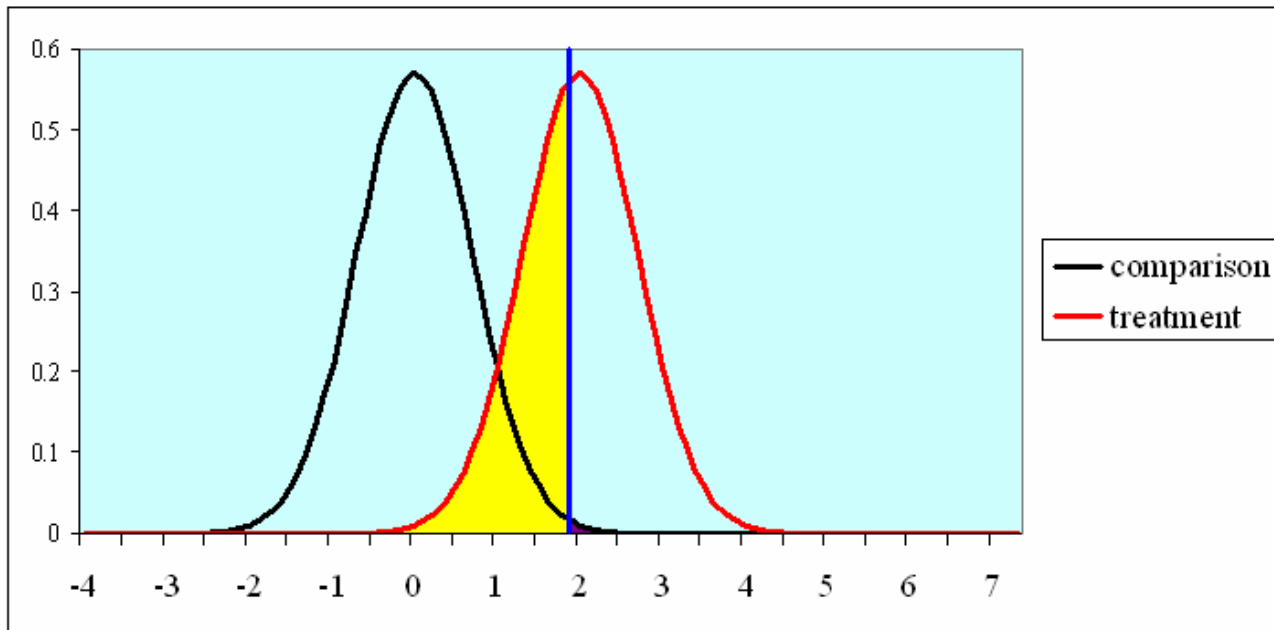
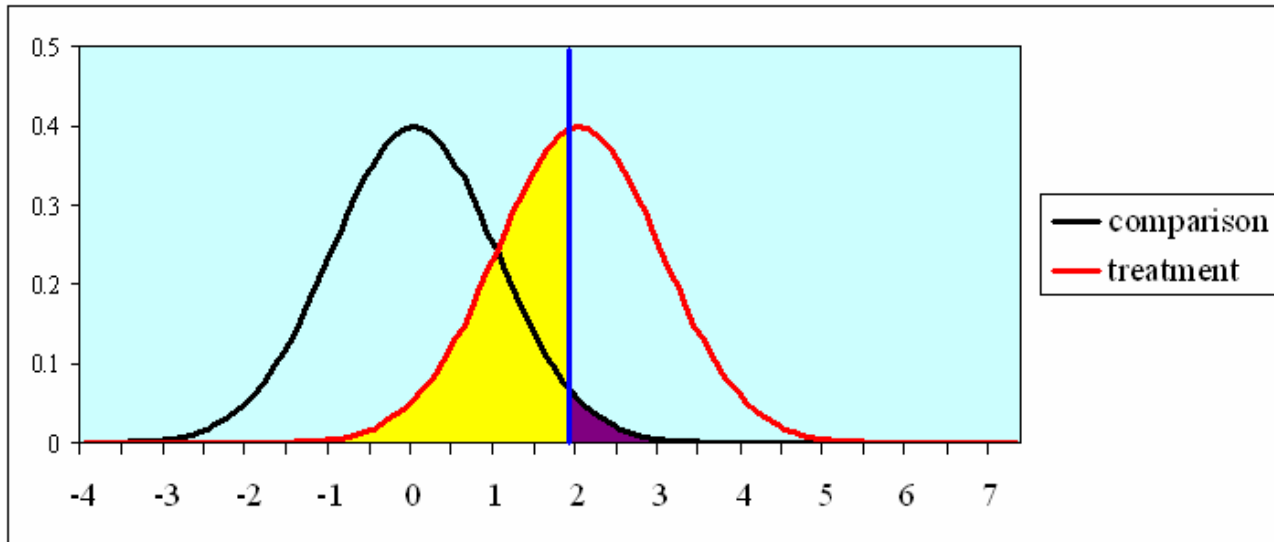
What happens when we increase the Type I error?



What happens when we increase the expected size of the effect?



What happens when we increase sample size?



تمرين منزلي: بعض من المحاكاة في المعطيات الإحصائية

* حساب حجم العينة

```
sampsi 130 135, alpha (0.05)
power(0.8) sd1(15) sd2(18) أحادية الجانب;
```

* لنقم بزيادة الحجم المتوقع للأثر \leq ستكفي عينة أصغر؛

```
sampsi 130 145, alpha(0.05) power(0.8)
sd1(15) sd2(18) أحادية الجانب;
```

* لنقم بزيادة القوة اللازمة \leq نحتاج إلى عينة أكبر؛

```
sampsi 130 135, alpha (0.05)
power(0.9) sd1(15) sd2(18) أحادية الجانب;
```

تمرين منزلي: المزيد من المحاكاة في المعطيات الإحصائية

* لنقم بحساب قوة اختبار؛

```
sampsi 130 135, alpha (0.05) sd1 (15)
sd2 (18) n1 (100) n2 (100) أحادية الجانب;
```

* لنقم بزيادة الحجم المتوقع للأثر == < تزداد القوة

```
sampsi 130 145, alpha (0.05) sd1 (15)
sd2 (18) n1 (100) n2 (100) أحادية الجانب;
```

* لنقم بزيادة حجم العينة == < تزداد القيمة؛

```
sampsi 130 135, alpha (0.05) sd1 (15)
sd2 (18) n1 (200) n2 (200) أحادية الجانب;
```

الجزء 2: أخذ العينات (← بيانات تعطي صورة صادقة)



□ مسح استقصائية نموذجية

■ الغرض: كسب المعرفة عن مجموعة عامة

□ مثال: دراسة قياس مستويات المعيشة/المسح القومي للأسر المعيشية

■ العيّنة: نموذج عن سكان البلد المعني

□ تقييم الأثر

■ الغرض: قياس تغيرات المؤشرات الرئيسية للمجموعة العامة التي يستهدفها الإجراء التدخلية

■ عملياً: قياس الفرق في المؤشرات بين المجموعات موضوع الدراسة ومجموعات الضبط

■ نأخذ عيّنة استراتيجية لكي تكون العيّنة نموذجية في المجموعات موضوع الدراسة ومجموعات الضبط

■ وهي ليست بالضرورة عيّنة نموذجية عن سكان البلد المعني

□ فقط إذا كان...

■ المسح العام للأسر المعيشية يعطي صورة صادقة عن مجموعات الدراسة ومجموعات المقارنة

■ هنالك مشاهدات كافية في مجموعات الدراسة ومجموعات الضبط

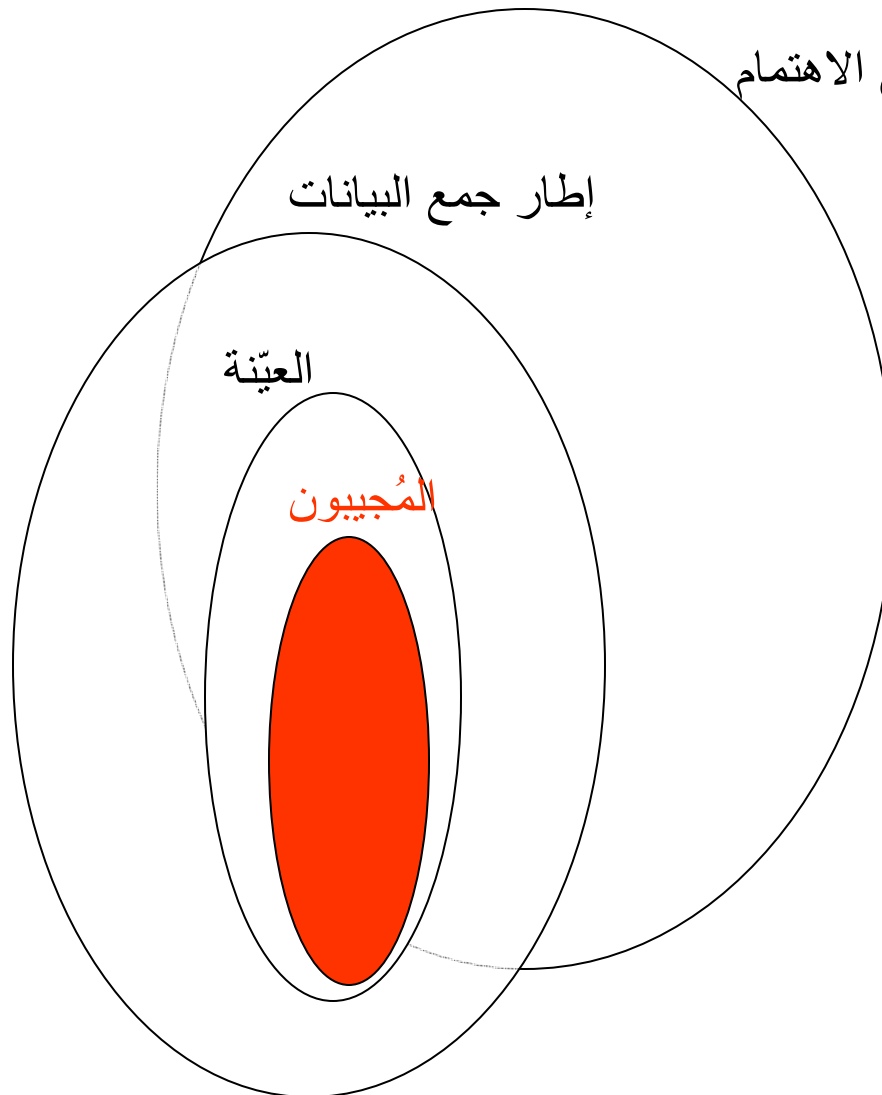
□ مثال: نُقيّم أثر حملة توعية الأمهات بشأن إرضاع أطفالهن دون السنة الأولى من العمر، 10 مناطق ضبط و 10 مناطق دراسة. هل يمكننا استخدام مسوحات الأوضاع الديموغرافية والصحية كأساس؟

■ الكثير من النساء في مسوحات الأوضاع الديموغرافية والصحية ليس لديهن أطفال دون السنة الأولى من العمر.

■ مسوحات الأوضاع الديموغرافية والصحية لا تعطي صورة صادقة عن الأوضاع على مستوى المنطقة المعنية.

- وحدة التحليل: نوع ما يُراد جمع المعلومات عنه (أشخاص، فُرى، مدارس ...)
- المجموعة العامة: مجموعة وحدات التحليل التي نود بالنسبة لها معرفة ما إذا نجحت الإجراءات التدخلية أم لا
- إطار جمع العينات: قائمة الوحدات التي ستؤخذ منها العينات (ما هو سيناريو أفضل حالة؟)
- العينة: وحدات التحليل التي نأخذها من إطار جمع البيانات، والتي نود جمع المعلومات عنها
- المُجيبون: وحدات التحليل التي من أجلها نحصل على البيانات (مثلاً: الأشخاص الذين يجيبون على أسئلة المسح الاستقصائي)
- نسبة الإجابات: النسبة المئوية لوحدات التحليل التي في العينة التي نجمع عنها المعلومات (أي الذين يُجيبون على أسئلة المسح الاستقصائي)

المجموعة العامة موضع الاهتمام



نسبة الإجابات ...

- سيناريو أفضل حالة: إنه 100 %
- حين تكون منخفضة، هنالك مخاطر تحييز الاختيار، على سبيل المثال حين لا تريد مجموعة الضبط الإجابة على أسئلة المسح الاستقصائي "تحييز عدم الإجابة"
- قواعد بديهية:
 - عدم تصديق النتائج إذا كانت نسبة الإجابات أقل من 70 في المائة بالنسبة لمجموعة الضبط أو مجموعة الدراسة
 - يجب عدم وجود فرق كبير في نسبة الإجابة بين مجموعة الضبط ومجموعة الدراسة
- لا بأس... ولكن ما الذي يحدث حين تكون نسبة الإجابات منخفضة؟

حين تكون نسبة الإجابات منخفضة ...

- ما الرأي بالنسبة للآتي: "زيادة حجم العينة سلفاً، لأننا نعرف أن بعض الناس لا يُجيبون"
- لا ينجح ذلك لأن أي تحييز عدم إجابة سيكون موجوداً أيضاً في العينة الأكبر حجماً
- أفضل الحلول:
- المزيد من الجهود في جمع المعلومات عن الذين لا يجيبون (أي 3 زيارات لبيوتهم بدلاً من زيارة واحدة)
- أخذ عينات فرعية من غير المجيبين لمعرفة ما إذا كانوا مختلفين عن الذين يُجيبون
- في مسح لاحق: النظر في الخصائص الأساسية للمجيبين وغير المجيبين لمعرفة ما إذا كانوا متماثلين
- دائماً: التأكد من أن الاستشاريين يُبلغون عن نسبة الإجابات في تقرير التقييم الذي يرفعونه

طرق أخذ العينات

□ العيّنات العشوائية

- بسيطة: لكل وحدة في إطار جمع العيّنات نفس احتمال الاختيار للاشتمال في العيّنة
- العيّنات الطبقيّة: أولاً نقسم إطار جمع العيّنات إلى طبقات (مجموعات)، ثم نأخذ عيّنة عشوائية بسيطة من كل طبقة
- العيّنات العنقودية: نأخذ مجموعات من الوحدات كعيّنة: مثال، قرى مع كافة من يعيشون فيها
- العيّنات متعددة المراحل – مزيج من طريقتين أو أكثر من أخذ العينات العشوائية من الطرق الأنفة، على سبيل المثال: عيّنة عشوائية طبقية من القرى ثم عيّنة عشوائية بسيطة من الأشخاص في القرى المشمولة في العيّنة

□ العيّنات المنهجية:

- مثال: دراسات الحالات: ينبغي توخي الحذر فمن غير الممكن تعميم نتائج هذه العيّنات واعتبارها بمثابة قوانين

أثر التصميم بالنسبة للعينات العنقودية/المتعددة المراحل



□ عند تجميع العينات في عينات عنقودية، يؤثر هذا في قوة اختبار الأهمية.

□ مثال: لنقم بتقدير طول قامة تلاميذ المدارس من سن 7 سنوات في مصر

■ الحالة 1: نقوم بمسح 200 تلميذ من سن سبع سنوات يتم اختيارهم من بين كافة تلاميذ المدارس المصرية

■ الحالة 2: أولاً، نختار عشوائياً 20 مدرسة من بين كافة مدارس مصر، ثم نختار عشوائياً 10 تلاميذ من كل من المدارس العشرين

أية عينة تعطينا أكبر قدر من المعلومات؟

العينات العنقودية تغير تباين المؤشرات

- الأشخاص ضمن مجموعة عنقودية أكثر تماثلاً مع بعضهم مما هم مع الأشخاص في مجموعة عنقودية أخرى:
- التلاميذ العشرون في المدرسة نفسها هم عادة من خلفيات أكثر تماثلاً من خلفيات 20 تلميذاً من مدارس مختلفة
- لذلك، الخطأ المعياري في متوسط طول قامة 200 تلميذ في 20 مدرسة (عينة طبقية وعشوائية) أكبر مما لو اخترنا 200 تلميذ عشوائياً (عينات عشوائية بسيطة)

$$Var_{Clustered} = Var_{SRS} * \underbrace{(1 + \rho * (k - 1))}_{\text{design effect}}$$

■ ρ هي مقياس التجانس بين الوحدات (الأطفال) المنتمية لنفس المجموعة العنقودية (المدرسة)

■ k = عدد الوحدات (الأطفال) المشمولين في عينة كل مجموعة عنقودية (مدرسة)

ماذا يعني هذا بالنسبة لحجم العينة

- التباين أكبر ← انخفاض القوة بالنسبة لحجم عينة محدد والخطأ من النوع 1(alpha)!
- فقدان القوة نتيجة أثر التصميم يمكن التعويض عنه بزيادة حجم العينات العامة
- إلى أي مدى يجب زيادة حجم العينة إذا تم اعتماد العينات العنقودية؟

تحديد العينة عند وجود عينات عنقودية

- الخطوة 1: ضرورة استخدام طرق معيارية لتحديد العينة بافتراض أخذ عينات عشوائية بسيطة (أنظر الجزء الأول من هذا العرض)
- الخطوة 2: ضرورة استخدام الصيغة التالية لتعديل حجم العينة من أجل العينات العنقودية

$$n_{Cluster} = n_{SRS} * \text{design effect}$$

$$= n_{SRS} * (1 + \rho * (k - 1))$$

- ينبغي تقدير ρ من مسوحات استقصائية قائمة ...
- ضرورة استخدام معطيات إحصائية!

متى نقلق فعلاً بشأن حجم العينة؟

□ عندما

- تكون العينات صغيرة جداً عند مستوى الوحدة موضوع الدراسة!
- نفترض مجموعة الدراسة في 20 مدرسة ومجموعة الضبط في 20 مدرسة

□ ولكن هنالك 400 تلميذ في كل مدرسة

■ هذه مازالت عينة صغيرة

□ عندما

- يكون للتقييم الذي نقوم به تصميم متداخل
- تزداد شروط حجم العينة زيادة كبيرة جداً

□ عندما

■ نستخدم عدم الاستمرارية التراجعية كإستراتيجية للتعليم

استنتاجات

- لسوء الحظ، في العديد من الحالات لا يمكننا استخدام مسوحات استقصائية عامة مثل دراسة قياس مستويات المعيشة ومسوح الأوضاع الديموغرافية والصحية كمسوح أساس.
- إطار جمع العينات وطريقة اختيارها يحددان مدى نموذجية العينة بشأن المجموعة العامة
- عندما نريد قياس فروق صغيرة جداً في المؤشرات فيما بين مجموعة الدراسة ومجموعة الضبط، نحتاج إلى عينة كبيرة جداً
- عندما تكون هنالك مجموعات عنقودية في العينة، نحتاج إلى عينة أكبر حجماً!