

3-3.Re: 統計小感想...

作者: yhliu (不務正業的老怪物) 看板: Statistics
標題: Re: 統計小感想...
時間: Tue Apr 6 20:24:08 1999

※ 引述《Alg (找我 請query)》之銘言：
： 嗯 系上派了二位老師 加強學生的經濟和統計
： 經濟 暫時不討論了...
： 那來討論統計好了
： 老師認為這學期是要讓學生跟著進度去念書
： 針對研究所考題或方向 是下學期的事了
： ^^

 嗯!
 所以...這學期是在奠定基礎觀念。

： 由學生分組上台針對各組負責的章節做討論
： 其實 一周一堂課 根本不夠
： 學生有很多問題 想問 確沒時間去問
 可能要利用下課時間再去問。
 上課時間總是很快就過去了...對一個教者、學者都用心的課程而言。

 這個版，或許可彌補上述問題...這就是當初開版的主要目的之一。

： 老師提出來的問題 是理論沒錯 但是 讓我產生了不小的疑惑
： 1.平均數容易受什麼影響? (極端值) ch1
 就 Q&A 而言，沒甚麼。
 但再深論：為甚麼平均數會受極端值影響?
 這有甚麼利、弊? 對策如何?

： 2.經驗法測是在否在任何情況下皆可用? 不可,要在鍾型分配下) ch1
 所謂「經驗法則」，當指1、2、3倍標準差範圍內之機率而言。
 為何「不可」? 為何在「鐘形分布」之下可? 又何時可認為「鐘
 形分布」條件成立? 細思之，了解其原由，而非只是 Q&A，則對
 解題、對實用，都有助益。

： 3.為甚麼探討機率分配時 到最後都在講隨機變數?
： (因為任何分配都可用隨機變數來表示) ch2
 仍是個需再思考的問題：
 何謂「機率分配」? 何謂「隨機變數」?
 為何討論「機率分配」? 隨機變數和實務有何相干?

： ：
： ：
： 諸如此類問題 讓我覺得 這些東西是否真的對未來準備研究所考試有用嗎?
 常常有要考研究所的網友來問問題。
 由所問問題，常覺得對方並未去仔細讀過一本教科書，
 而是拿補習班所出問題集代替。
 就「準備研究所」這個層次而言，基礎觀念的了解是否
 有幫助，或許見仁見智。我個人是這麼走過來的...不過，
 我學的是統計，和學企管的情形當然不同。企管所需要的
 的統計知識，要到甚麼程度;目前你們上課的內容，是
 否直接對提昇考試成績有幫助，不敢置評。不過，我很
 敬佩貴系老師願意花時間補強你們的基礎。這基礎即使
 對考試沒有明顯功效，對未來在研究路上，了解所需統

stlearn-3.txt

計方法，我想多少是有些幫助的。不過，重要的是去了解。如果只是記誦 Q&A，如上文所列形式（我想那並不表示你所著重只是 Q&A），則它可能一點用處都沒有...

: 在摸不清頭緒下 也只能跟著進度去走
: 在沒有學長姊可詢問下 這條路想必會走的很艱苦...
就準備考試而言，看看過去的考古題，是有幫助的。
對觀念的解、討論，希望本版能發揮一些功效! :)
感冒中，有點不曉得自己在打什麼的Alg

嗨! guest你好! 今天過得好嗎? 一定 happy 對不對? :)
祝你天天 happy 啦! :)

交大資科站 (140.113.23.3, bbs.cis.nctu.edu.tw) 機率與統計 (Mprobability) 版
本站 (140.116.96.199, MoonStar.twbbs.org) 統計: 讓數字說話 (Statistics) 版
竭誠期待你的光臨喔! ^ ^ :)
我是老怪物...瘋狂老怪物...我最怕你來打聽我的底細...所以...我要溜了! :p

--
※ Origin: 盈月與繁星 (MoonStar.twbbs.org) ◆ From: 140.116.52.147

> ----- <

作者: yhliu (不務正業的老怪物) 看板: Statistics
標題: Re: 統計小感想...
時間: Tue Apr 6 22:48:21 1999

※ 引述《Alg (找我 請query)》之銘言：
: ※ 引述《yhliu (不務正業的老怪物)》之銘言：
: : 仍是個需再思考的問題：
: : 何謂「機率分配」？何謂「隨機變數」？
: : 為何討論「機率分配」？隨機變數和實務有何相干？
: : 在我的想法裡
: : 1.隨機變數的存在最大的目的 在於表達樣本的組合不是數字時 方便處理
: : 2.致於任何分配都可用隨機變數來表示....在我的看法裡 是對沒錯
: : 但到底隨機變數的存在的因素 是1>2? or 2>1?
如果我們把「機率分配」看成是
「數值性隨機現象的機率配置法則」，
那麼，問題可能就較清楚了! :)

在我們所感興趣的現象，有很多是數值性的，例如擲骰子的點數，一批產品中的不良品個數，一群人中對某項產品滿意的人數等等。有些現象可能不是數值性的，如產品的良、不良，如撲克牌的點數和花色等等。但這些現象，我們可透過「數量化」使它比較方便處理，如前述良品、不良品轉成「不良品個數」；撲克牌花色點數依目的轉成點數 (1~13) 或 (花色,點數) 的數對；如市場研究 (行銷研究) 中消費者偏好改成適當量化指標等。

把我們觀測的隨機現象 (如隨機抽樣問消費者的偏好，做工業實驗觀測產量或產品品質) 做數量化處理，成為數值性隨機現象。這樣的對應關係，就是「隨機變數」。這裡有一個重要觀念，在初等統計中可能沒講清楚的，就是「隨機變數」事實上是一個函數：由實際觀測結果映至一個數值的函數關係。例如研究產品的良品、不良品，觀測現象是每一被觀測到的產品判斷結果 (良、不良)。但我們所關心的只是究竟不良率會不會太高？因此，我們將這個觀測結果映至一個數值：就是這批被觀測

的樣本中，有多少是不良品？

顯然：討論「有多少不良品」比討論「每一個樣品是良品或不良品」來得簡單！

而，「機率分配」廣義來說雖然可代表任何隨機現象中各事件（或各結果）的機率；但狹義而言，常僅代表數值性隨機現象各事件的機率。如：一批 $n=100$ 個樣品中有 x 個是不良品的機率。

如果不用「隨機變數」，那麼，我們要看的「機率分配」可能是： n 個產品（假設按某種方式編號）結果是（良，壞，壞，良，...良）的機率是多少？

用隨機變數，我們所討論的問題較簡略，只關心： n 個產品，有 x 個不良的機率是多少？

因此，你的第一點了解是完全正確的！使用隨機變數的目的，就是要簡化觀測結果。而就第二點，措辭上是有點怪怪的...不能完全理解你的意思。如果你的意思是說：我們所觀測的現象可以做適當量化，那是對的！：)

： 致於你說的機率分配 我倒有下面的想法
： 一般應用時，並不知隨機變數所對應之分配，因此統計學者的工作就是對這些
： 些隨機變數建立機率模式。感覺上，統計學在建立這些模式很有可信度，當然
確實，統計學雖建立在機率論的基礎上，但和機率論所談的有很大的
不同，那就是：統計人員並不假設我們知道所觀測現象（用隨機變數
做量化後）的機率分配。如果這個機率分配已確知了，那就不用不著統
計人員了！統計人員的基本目標，就是去「猜測」這個機率分配。：p
猜測的方法，可能在很廣泛，無特定形式的基礎上做猜測（無母數方
法）；也可能限制在很小的範圍，例如假設是指數分配、常態分配、二
項分配、卜瓦松分配等，只是有少數「未知參數」需猜測（有母數方
法）。統計人員可直接猜測一個結果（點估計）、一個範圍（區間估計），
或在兩個範圍中做選擇（假說檢定）等。而統計人員做這些猜測，是利
用資料中會透露出正確機率分配的訊息。因此，可以有適當的可信度。

： 前面文章說的 統計是一門騙人的科學....倒令我有一些感悟.....
： 統計是一門說服人的科學或學科
不錯！：)

： 即使alpha值界於臨界值 是否可以用人為的操作去得出自己要得的結果？
： （這好像是會計選擇不同的會計處理方法，去做出最令大眾滿意的盈餘
： 即使在會計學上 變更會計處理方法都要做附註，但是統計學就....）
： 公眾的alpha值通常是0.05 確實要反駁不易..的確是自由心證了...
顯著水準或信賴水準的決定，統計學上並未提供明確的準則，
而一般常用 5% 或 1%，只是相因成習的做法。如果追究到顯
著水準的原始意義，可以了解為何統計學家不提供一個標準，
以及為何大家習慣用 5% 或 1%。因為顯著水準代表我們所願
容忍的某種錯誤（檢定上的型I錯誤）的機率。這個錯誤機率
可容忍多少是因問題、因目的而不同的，因為它代表一種風險...
但此風險還和實際發生錯誤時可能招致的損失有關。而錯誤並
不只一種；因此，還要看另一種錯誤發生時的損失，才能決定
多大的顯著水準為適當的。但對於一些目的不是很明確、有多
重目的的統計分析結果而言（如科學研究），要明確說出不同錯
誤造成多少損失，有時卻很難。而 5% 或 1% 的錯誤機會，在
一般狀況，一般人可能認為「已經夠小了」，因此願意接受這
種標準。但這個標準，當然也會因不同目的而有所調整。當型I

錯誤是很嚴重的錯誤時，或許我們會採取 0.1% 的標準。

至於統計方法及標準的揭露，和會計方法變更並不相同，但它也是必要的！在正式的統計報告，必須揭露資料蒐集方法（如抽樣設計、調查或測量方式、變數定義等等）、統計模型（假設條件、機率模型、數量關聯模型等等）、統計方法（估計方法、檢定統計量及決策程序等等）及誤差估計（估計標準誤、顯著水準、信賴度等）。

--
嗨! guest你好! 今天過得好嗎? 一定 happy 對不對? :)
祝你天天 happy 啦! :)
交大資科站 (140.113.23.3, bbs.cis.nctu.edu.tw) 機率與統計 (Mprobability) 版
本站 (140.116.96.199, MoonStar.twbbs.org) 統計:讓數字說話 (Statistics) 版
竭誠期待你的光臨喔! ^ ^ :)
我是老怪物...瘋狂老怪物...我最怕你來打聽我的底細...所以...我要溜了! :p

--
※ Origin: 盈月與繁星 (MoonStar.twbbs.org) ◆ From: 140.116.52.147

> ----- <

作者: yhliu (不務正業的老怪物) 看板: Statistics
標題: Re: 統計小感想...
時間: Tue Apr 6 23:21:37 1999

※ 引述《Alg (找我 請query)》之銘言：
: ※ 引述《yhliu (不務正業的老怪物)》之銘言：
: : 因此，你的第一點了解是完全正確的！使用隨機變數的目的，就是
: : 要簡化觀測結果。而就第二點，措辭上是有點怪怪的...不能完全理
: : 解你的意思。如果你的意思是說：我們所觀測的現象可以做適當量
: : 化，那是對的！ :)
: 第二點是老師說的....
: 他的問題是問 為什麼討論到機率時 到最後都在討論隨機變數
: 我回答第一點.....他說也沒錯啦
: 可是重點應該是在第二點 任何情況下都可以量化
: 但我一直在思考 到底是他的講法對 還是我的講法對
: 還是說 我第一點回答的 只是 隨機變數存在的意義
: 而第二點回答的是 為什麼討論到機率時 到最後都在討論隨機變數
: 嗯 好像越討論越清楚了 ^_^
: ^^^^^^^^^^^^^^^^^^^

這樣很好，不是嗎? :)

說「任何情況都可以量化」也許是的，但問題是：何謂「量化」?

如前面討論，「可以量化」應是一切討論的基礎；但「為何量化」才是討論隨機變數（及其機率分配）的重點。我的意見是：兩者如陰陽互生，缺一不可！因此，我不會去爭論何者較重要。 :)

如前所述，量化的目的是我們將隨機現象量化的原動力。如果我們要的不只是一批產品的良率/不良率，而是要找出所有不良品予以置換，這便沒有量化問題了！在消費者行為研究，我們可能興趣在消費者對某項產品的綜合評價，則我們可能建立一指標，由被查者在多個相關問項的回應對應出一指標值，這是一種量化方式。但另

stlearn-3.txt

一研究或許關心的是消費者回應類型，則量化的方式是對每一問項做量化，構成回應向量。而若問項的回答並無順序性（如「喜歡的顏色」），在前一研究中即很難量化（不能說喜歡紅色的比喜歡藍色的應給予較高分數）；而在後一研究，則一個問項需以一個向量做量化。