BỘ GIÁO DỤC VÀ ĐÀO TẠO
DỰ ÁN ĐÀO TẠO GIÁO VIÊN THCS
LOAN No 1718 - VIE (SF)

NGUYỄN ĐÌNH HIỀN

Giáo trình
XÁC SUẤT THỐNG KÊ

$$p(A_i / B) = \frac{p(A_i).p(B/A_i)}{p(B)} \quad (i = \overline{1,n})$$

$$p(A_i / B) = \frac{p(A_i).p(B/A_i)}{p(B)} \quad (i = \overline{1,n})$$

SP

NHÀ XUẤT BẢN ĐẠI HỌC SƯ PHẠM

NGUYỄN ĐÌNH HIỀN

Giáo trình
XÁC SUẤT THỐNG KÊ

(Giáo trình Cao đẳng Sư phạm)

NHÀ XUẤT BẢN ĐẠI HỌC SƯ PHẠM

Mã số: 01.01. 23/191. ĐH. 2006

MỞ ĐẦU

Xác suất thống kê là một ngành khoa học được dạy trong các trường Đại học và Cao đẳng của gần như tất cả các ngành, kể cả tự nhiên và xã hội, tuy nhiên nội dung dạy có khác nhau. Tuỳ yêu cầu của từng ngành mà chỉ định số tiết, trong các ngành kĩ thuật sinh học và nông nghiệp thường dạy từ 45 đến 75 tiết, nội dung cũng được lựa chọn khác nhau.

Giáo trình Xác suất thống kê này được viết cho sinh viên Cao đẳng Sư phạm Kĩ thuật Nông nghiệp. Nội dung dựa trên chương trình Xác suất thống kê khối B của Bộ Giáo dục và Đào tạo nhưng viết lại theo khung chương trình đào tạo Cao đẳng Sư phạm Kĩ thuật Nông nghiệp cho phù hợp với thời lượng và yêu cầu.

Giáo trình cố gắng cung cấp cho học viên một số kiến thức cơ bản về Xác suất và thống kê để có cách nhìn biện chứng hơn về các hiện tượng tự nhiên và xã hội, để hiểu kĩ hơn một số phần mang tính định lượng trong sinh học và có cơ sở để học môn Phương pháp thí nghiệm nên chỉ trình bày một cách đơn giản các khái niệm xác suất và biến ngẫu nhiên, kèm theo nhiều thí dụ minh hoạ. Phần thống kê chỉ trình bày kĩ mục đích của từng vấn đề, các bước tính, cách kết luận và các thí dụ minh hoạ.

Để nắm được kiến thức trình bày trong sách không có cách nào tốt hơn là xem kĩ thí dụ và làm đầy đủ bài tập.

Giáo trình viết cho người học, do đó khi dạy các giáo viên cần tham khảo thêm các sách viết kĩ hơn, sâu hơn về Xác suất thống kê toán học như các giáo trình dùng cho khối sinh của Đại học Quốc gia, Đại học Sư phạm hay Đại học Nông nghiệp.

Phần bài tập có bài giải mẫu và đáp số. Vì học viên đã quen với tin học nên giáo trình cung cấp thêm một số chương trình đơn giản viết bằng ngôn ngữ Pascal để học viên có thể tự mình tính toán các bài tập xác suất thống kê và chuẩn bị cho sau này học môn Phương pháp thí nghiệm.

Trong giáo trình các phần đánh dấu * có thể bỏ qua, nếu có điều kiện thì đọc để mở rộng kiến thức.

Sau đây là nội dung chính của giáo trình:

Chương 1 trình bày các khái niệm cơ bản trong giải tích tổ hợp, nếu học viên đã học rồi (phần này hiện đã dạy ở nhiều trường Trung học phổ thông) thì chỉ nhắc lại và củng cố qua bài tập.

Chương 2 trình bày các khái niệm cơ bản về Xác suất, đây là chương quan trọng và rất khó dạy, do đó phải khéo léo kết hợp giữa cách trình bày sao cho không trừu tượng quá mà vẫn đảm bảo tính chặt chẽ, vì thực chất chương này chính là hệ tiên đề của môn Xác suất. Yêu cầu cần đạt được là giới thiệu mô hình suy luận sau: Phép thử có các kết quả trực tiếp, gọi là các sự kiện sơ cấp, sự kiện là tập hợp một số sự kiện sơ cấp, xác suất là số đánh giá khả năng xuất hiện của sự kiện. Xác suất tuân theo một số quy tắc tính và yêu cầu phải nắm được hai quy tắc cộng và nhân tổng quát và đơn giản.

Chương 3 giới thiệu khái niệm biến ngẫu nhiên, phần này không nên sa vào các định nghĩa trừu tượng mà phải thật cụ thể, do đó cần theo dõi các thí dụ, qua đó tổng hợp nên khái niệm biến ngẫu nhiên, bảng phân phối, hàm phân phối. Phần số đặc trưng có thể dạy sơ qua, chú ý đến ý nghĩa của kì vọng và phương sai chứ không đi sâu chứng minh các tính chất.

Chương 4 cần trình bày kĩ phân phối nhị thức và phân phối siêu bội. Trong phần biến liên tục chỉ tập trung trình bày phân phối chuẩn và cách tính gần đúng phân phối nhị thức bằng phân phối chuẩn.

Với thời lượng 15 tiết, phần này không nên học hoặc dạy tràn lan mà chỉ tập trung vào một số điểm chính, tuy nhiên giáo trình vẫn viết đầy đủ để học viên tham khảo. Phần bài tập đã chọn các bài phù hợp với trình độ cao đẳng, không khó quá, nhưng cũng không thể coi là quá dễ.

Phần thống kê bắt đầu bằng ***chương 5***, giới thiệu khái niệm tổng thể, mẫu quan sát và các tham số của mẫu quan sát, tiếp theo là công thức ước lượng trung bình μ của biến phân phối chuẩn và xác suất p của phân phối nhị thức. Chương này không yêu cầu trình bày lí thuyết mà phải thật cụ thể, học xong phải biết cách tính trung bình cộng, phương sai mẫu, cách tra cứu bảng $\varphi(u)$, $\Phi(t)$, t và biết cách ước lượng μ, p.

Chương 6 cũng chỉ trình bày rất ngắn gọn bài toán kiểm định giả thiết, giả thiết và đối thiết, giới thiệu quy tắc kiểm định giá trị trung bình của một biến phân phối chuẩn và bài toán so sánh hai trung bình của hai tổng thể phân phối chuẩn. Chương này để tiết kiệm thời gian có thể trình bày bằng bảng kẻ sẵn, nêu các trường hợp gặp phải khi kiểm định, công thức tính, cách kết luận (tương tự như ở phụ chương 2).

Chương 7 trình bày kiểm định một phân phối và bảng tương liên. Cả hai phần này liên quan đến biến định tính và đều dùng phân phối Khi bình phương

(χ^2) do đó khi trình bày cũng có thể dùng bảng kẻ sẵn để làm nổi bật nội dung và cách làm rất giống nhau của hai phần (xem phụ chương 2).

Chương 8 giới thiệu tương quan và hồi quy tuyến tính, nếu ít thời gian thì chỉ trình bày ý nghĩa hệ số tương quan, cách tính, các kết luận. Phần hồi quy tuyến tính chỉ trình bày ý nghĩa của mô hình tuyến tính để tính xấp xỉ biến ngẫu nhiên Y theo biến đã cho X, cách tính các hệ số, kết luận.

Phần đáp số trình bày gần hết các đáp số của các bài tập của các chương, kể cả bài tập thường và bài tập có ghi dấu *

Phụ chương 1 giới thiệu một số chương trình viết bằng ngôn ngữ Pascal dưới dạng thật đơn giản để học sinh, nếu đã học tin học và có điều kiện sử dụng máy tính, có thể tự mình tính toán xác suất và thống kê trên máy tính cũng như tự tạo ra bảng tính để tra cứu.

Phần phụ chương 2 có bảng ghi các thuật ngữ xác suất thống kê dùng trong giáo trình và các công thức. Phần công thức có thể dùng để tham khảo khi trình bày phần thống kê sao cho ngắn gọn, dễ hiểu.

Cuối cùng là các bảng tính, các bảng này rất cần cho phần thống kê nên khi dạy phải chỉ cho học viên cách tra cứu cả xuôi lẫn ngược.

Giáo trình đã nhận được sự góp ý chân tình, chính xác và tỉ mỉ của Phó giáo sư, Tiến sĩ Đào Hữu Hồ và Phó giáo sư, Tiến sĩ Tô Cẩm Tú. Tác giả xin chân thành cảm ơn.

Viết giáo trình là việc khó và càng khó khi thời lượng tương ứng của môn học lại rất ít. Chắc chắn cuốn sách này còn nhiều thiếu sót, rất mong sự góp ý của bạn đọc. Xin chân thành cảm ơn.

Hà Nội, tháng 1 năm 2003

Tác giả

Chương 1

GIẢI TÍCH TỔ HỢP

Chương này không nằm trong nội dung môn Xác suất thống kê mà thuộc về các kiến thức chung đã được dạy ở Phổ thông, tuy nhiên để hiểu được các phép tính xác suất, thống kê ở các chương sau thì cần phải học, hoặc nếu đã học rồi thì ôn lại các khái niệm cơ bản như: chỉnh hợp, hoán vị, tổ hợp, chỉnh hợp lặp, nhị thức Niu-tơn.

§1. CHỈNH HỢP

Thí dụ 1

Cửa hàng có 3 cái mũ màu xanh, đỏ, tím. Có 2 khách đến mua, cô bán hàng lấy lần lượt ra 2 cái mũ giao cho 2 khách, cái thứ nhất màu xanh, cái thứ hai màu đỏ, ta kí hiệu tắt kết quả này là (X, Đ), cũng có thể cái thứ nhất màu đỏ, cái thứ hai màu xanh (Đ, X), hoặc (X, T), (T, X), (Đ, T), (T, Đ). Ta gọi mỗi kết quả là một chỉnh hợp chập 2 trong 3 vật, có tất cả 6 chỉnh hợp chập 2 của 3 mũ. Có thể lập luận như sau: Cái mũ chọn đầu tiên là bất cứ mũ nào trong 3 mũ, như vậy có 3 cách chọn, sau đó có 2 cách chọn mũ thứ hai, như vậy có 3.2 = 6 cách chọn lần lượt 2 trong 3 mũ. Hai cách chọn (X, Đ) và (X, T) khác nhau vì có một mũ khác nhau, còn 2 cách chọn (X, Đ) và (Đ, X) thì khác nhau về thứ tự chọn.

Thí dụ 2

Một tổ có 10 người, chọn lần lượt 3 người đi làm việc, người thứ nhất là nhóm trưởng, người thứ hai theo dõi các chỉ tiêu kinh tế, người thứ ba theo dõi các chỉ tiêu kĩ thuật. Giả sử 10 người trong tổ có khả năng làm việc như nhau thì có 10 cách chọn nhóm trưởng, sau đó có 9 cách chọn người phụ trách chỉ tiêu kinh tế và cuối cùng có 8 cách chọn người thứ ba. Gọi mỗi nhóm 3 người như vậy là một chỉnh hợp chập 3 của 10 người có tất cả 10.9.8 = 720 chỉnh hợp chập 3 của 10 người.

Hai nhóm khác nhau nếu có ít nhất một thành viên khác nhau hoặc thành viên của nhóm giống nhau nhưng thứ tự chọn khác nhau, do đó phân công công việc trong nhóm khác nhau.

Thí dụ 3

Có 8 đội bóng chuyền vào chung kết. Có 3 đội sẽ được huy chương: một đội được huy chương vàng, một đội được huy chương bạc, một đội được huy chương đồng. Nếu 8 đội thực lực như nhau thì có thể có bao nhiêu dự báo về danh sách bộ ba được huy chương? Ta lại lập luận như ở thí dụ 2, vì thực lực như nhau nên có thể có 8 cách dự báo đội được huy chương vàng, sau đó còn 7 cách dự báo đội được huy chương bạc, cuối cùng có 6 cách dự báo đội được huy chương đồng, như vậy tất cả có 8.7.6 = 336 chỉnh hợp chập 3 của 8 đội. Hai dự báo khác nhau nếu trong danh sách 3 đội được huy chương có ít nhất tên một đội khác nhau hoặc vẫn cùng tên 3 đội nhưng thứ tự khác nhau do đó có sự thay đổi tên đội tương ứng với loại huy chương.

Tổng quát. Có n vật khác nhau lấy lần lượt ra k vật, mỗi nhóm k vật như vậy được gọi là một chỉnh hợp chập k của n vật. Nếu vật nào cũng có khả năng được chọn như nhau thì có n cách chọn vật thứ nhất, (n – 1) cách chọn vật thứ hai, ..., (n – k + 1) cách chọn vật thứ k. Tất cả có n(n – 1) ... (n – k + 1) chỉnh hợp chập k của n vật. Hai chỉnh hợp khác nhau nếu có ít nhất một vật khác nhau hoặc vật như nhau nhưng thứ tự lấy ra khác nhau.

Định nghĩa. *Một nhóm k vật lấy lần lượt trong số n vật khác nhau gọi là một chỉnh hợp chập k của n vật. Số chỉnh hợp chập k của n vật, kí hiệu là* A_n^k, *được tính theo công thức:*

$$A_n^k = n(n-1)...(n-k+1) \quad (1 \le k \le n) \qquad (1.1)$$

§2. HOÁN VỊ

Thí dụ 4

Trong thí dụ 1 có 3 khách đến mua mũ, giả sử cô bán hàng lấy cả 3 mũ và đưa lần lượt cho 3 khách, nếu khách thứ nhất nhận mũ xanh, khách thứ hai nhận mũ đỏ, khách thứ ba nhận mũ tím thì ta có kết quả (X, Đ, T), nhưng có thể cô bán hàng chọn mũ theo thứ tự khác nên kết quả là (Đ, X, T) hay (T, Đ, X), ... tất cả có 6 kết quả khác nhau.

Vì có 3 mũ lấy cả 3 nên hai kết quả chỉ khác về thứ tự đưa 3 mũ cho 3 khách hàng, chẳng khác nào để 3 mũ X, Đ, T bên cạnh nhau sau đó đổi chỗ (hoán vị) các mũ, sau mỗi lần đổi chỗ được một kết quả khác, do đó mỗi kết quả gọi là một hoán vị của 3 mũ. Nếu nói theo cách trình bày ở thí dụ 1.1 thì mỗi hoán vị ở đây chính là một chỉnh hợp chập 3 của 3 mũ.

Thí dụ 5

Có 4 người bạn A, B, C, D đi xem văn nghệ và chọn 4 ghế ngồi cạnh nhau, nếu sắp A ngồi vào ghế 1, B ngồi ghế 2, C ngồi ghế 3, D ngồi ghế 4 thì có một cách sắp xếp 4 người vào 4 chỗ. Nếu đổi chỗ 2 người thì được một cách sắp xếp mới, mỗi cách sắp xếp như vậy gọi là một hoán vị.

Nếu nhìn theo góc độ chỉnh hợp thì có 4 người lần lượt chọn cả 4 và thứ tự chọn chính là số ghế, như vậy mỗi hoán vị chính là một chỉnh hợp chập 4 của 4 người, dùng công thức (1.1) có số hoán vị của 4 người là 4! = 4.3.2.1 = 24.

Thí dụ 6

Có 6 cụ ông sắp hàng ngang để tập thể dục buổi sáng, sau buổi tập đầy phấn khích các cụ quyết định từ ngày hôm sau sẽ ra tập tiếp và mỗi ngày sẽ sắp hàng theo một trật tự khác những lần tập trước. Hỏi sau bao nhiêu ngày các cụ mới quay lại cách xếp hàng đầu tiên?

Coi mỗi cách sắp hàng là một cách sắp xếp 6 cụ vào 6 chỗ, tức là một hoán vị của 6 cụ, cũng có thể coi đó là một chỉnh hợp chập 6 của 6 cụ, có thể tính được tất cả có 6! = 720 cách xếp hàng. Như vậy phải 720 ngày sau, tức là gần 2 năm sau 6 cụ mới xếp hàng lại theo đúng cách sắp hàng đầu tiên.

Tổng quát. Có n vật khác nhau được sắp xếp vào n chỗ, có n cách chọn vật thứ nhất để xếp vào chỗ thứ nhất, (n – 1) cách chọn vật thứ hai vào chỗ thứ hai, ... , (n – k + 1) cách chọn vật thứ k để sắp vào chỗ thứ k... Mỗi cách sắp xếp được gọi là một hoán vị của n vật.

Định nghĩa. *Một nhóm n vật được sắp xếp vào n chỗ, mỗi cách sắp xếp được gọi là một hoán vị. Mỗi hoán vị là một chỉnh hợp chập n của n vật.*

Số hoán vị được tính theo công thức:

$$A_n^n = n(n-1)\ ...\ 3.2.1 = n! \qquad (1.2)$$

§3. TỔ HỢP

Thí dụ 7

Trong thí dụ 1, cô bán hàng chọn 2 trong 3 mũ, có thể có 3 cách chọn: một xanh một đỏ, một xanh một tím, một đỏ một tím. Gọi mỗi cách là một tổ hợp chập 2 của 3 mũ. Sau khi chọn xong thì có 2 cách đưa cho hai khách tức là có 2 chỉnh hợp chập 2. Thí dụ chọn tổ hợp (X, Đ) thì có thể đưa mũ xanh cho khách thứ nhất, đưa mũ đỏ cho khách thứ hai hay đổi chỗ (hoán vị) hai mũ, đưa mũ đỏ cho khách thứ nhất đưa mũ xanh cho khách thứ hai. Như vậy ta có hệ thức:

3 (tổ hợp chập 2 của 3 mũ) × 2 (hoán vị của 2 mũ) = 6 (chỉnh hợp chập 2 của 3 mũ).