

Bài Tập Thông kê

1 Chương 1: Ước lượng

Bài 1. Tiến hành đo chiều cao của 100 học sinh lớp 3 ở một số trường trung tiểu học ở một huyện, ta thu được kết quả như sau

Khoảng chiều cao (cm)	số em
[110; 112)	5
[112; 114)	8
[114; 116)	14
[116; 118)	17
[118; 120)	20
[120; 122)	16
[122; 124)	10
[124; 126)	6
[126; 128)	4

Tính trung bình mẫu và phương sai mẫu.

Bài 2. Có tài liệu về lượng nước tiêu thụ (m^3 /tháng) của 100 hộ gia đình tại huyện X như sau:

Lượng nước tiêu thụ (m^3 /tháng)	Tần số
[0; 25)	12
[25; 50)	33
[50; 75)	40
[75; 100)	10
[100; 125)	5

Yêu cầu:

1. Tính lượng nước tiêu thụ trung bình của các hộ gia đình trong một tháng trong mẫu trên?
2. Tính độ lệch của lượng nước tiêu thụ của các hộ gia đình trong 1 tháng trong một tháng trong mẫu trên?

Bài 3. Công ty bao bì Hải Pack đang nhập lô hàng 20000 bao hạt nhựa của một nhà cung cấp quen. Dữ liệu quá khứ cho thấy khối lượng của các bao hạt nhựa này tuân theo luật phân phối chuẩn với phương sai $36(kg^2)$. Chọn ngẫu nhiên 25 bao hạt nhựa để cân thu được giá trị trung bình là 96 Kg/bao Với độ tin cậy 95% hãy ước lượng khoảng tin cậy của khối lượng trung bình của 2000 bao hạt nhựa này.

Bài 4. Doanh số của một cửa hàng là đại lượng ngẫu nhiên phân phối chuẩn với độ lệch chuẩn là 2 triệu đồng/tháng. Điều tra ngẫu nhiên doanh số của 600 cửa hàng có quy mô tương

tự nhau tìm được doanh số trung bình là 8,5 triệu. Với độ tin cậy 95% hãy ước lượng doanh số trung bình của các cửa hàng thuộc quy mô đó.

Bài 5. Để nghiên cứu nhiệt độ trung bình trong tháng 4 ở thành phố A, người ta theo dõi trong 10 địa điểm và thu được số liệu sau:

24,1	27,2	26,7	23,6	26,4
25,8	27,3	23,2	26,9	27,1

Với độ tin cậy 95% hãy ước lượng khoảng nhiệt độ trung bình trong tháng 4 của thành phố trên. Biết nhiệt độ trung bình trong 1 tháng là đại lượng ngẫu nhiên phân phối chuẩn.

Bài 6. Với độ tin cậy 95%, hãy ước lượng khoảng lượng xăng hao phí trung bình cho một ô tô chạy từ A đến B nếu chạy thử 40 lần trên đoạn đường này người ta ghi nhận được lượng xăng hao phí như sau:

Lượng xăng hao phí (lít)	Tần số
[9, 6; 9, 8)	3
[9, 8; 10, 0)	8
[10, 0; 10, 2)	10
[10, 2; 10, 4)	11
[10, 4; 10, 6)	8

Bài 7. Để định mức thời gian gia công một chi tiết máy, người ta theo dõi ngẫu nhiên quá trình gia công 35 chi tiết và thu được số liệu sau:

Thời gian gia công (phút)	Tần số
[15; 17)	1
[17; 19)	3
[19; 21)	9
[21; 23)	12
[23; 25)	8
[25; 27)	2

Bằng khoảng tin cậy đối xứng hãy ước lượng khoảng thời gian gia công trung bình một chi tiết máy với độ tin cậy $1 - \alpha = 0,95$.

Bài 8. Hãy ước lượng tỷ lệ chính phẩm của một nhà máy bằng khoảng tin cậy đối xứng với độ tin cậy 0,95 biết rằng kiểm tra 100 sản phẩm của nhà máy thì thấy có 10 phế phẩm.

Bài 9. Mở 200 hộp của một kho đồ hộp, người ta thấy có 28 hộp bị biến chất. Với độ tin cậy 0,95, bằng khoảng tin cậy đối xứng, hãy ước lượng tỷ lệ đồ hộp biến chất ở trong kho.

Bài 10. Trong đợt vận động bầu cử tổng thống người ta phỏng vấn ngẫu nhiên 1600 cử tri thì được biết 960 người trong số đó sẽ bỏ phiếu cho ứng cử viên A. Với độ tin cậy 90%, ứng cử viên A sẽ chiếm được tỷ lệ phiếu bầu trong khoảng nào?

Bài 11. Nhà máy A sản xuất 1 loại sản phẩm. Để ước lượng tỉ lệ thành phẩm người ta chọn ngẫu nhiên 400 sản phẩm và chia thành 40 nhóm để kiểm tra. Kết quả thu được như sau

Số thành phẩm trong nhóm	1	2	3	4	5	6	7	8	9	10
Số nhóm	2	1	3	6	8	10	4	5	1	0

Với độ tin cậy 90% hãy ước lượng khoảng tỉ lệ thành phẩm của nhà máy.

Bài 12. Để ước lượng số cá trong một hồ nuôi cá, người ta đánh bắt 3000 con, đánh dấu rồi thả xuống. Vài ngày sau lại đánh bắt được 500 con cá thấy 60 con cá được đánh dấu. Với độ tin cậy 95 % số cá trong hồ có bao nhiêu con?

Bài 13. Một nhà máy đưa ra thị trường một loại sản phẩm mới. Để xem đánh giá của người tiêu dùng đối với loại sản phẩm này như thế nào, người ta phát cho mỗi người mua hàng một phiếu thăm dò và yêu cầu gửi lại cho xí nghiệp sau 3 tháng. Kết quả sau 3 tháng người ta nhận được 300 phiếu thăm dò trong đó có 90 phiếu tỏ ra thích loại sản phẩm này. Hãy ước lượng tỷ lệ thực khách hàng thích loại sản phẩm này. Với độ tin cậy 95% tỷ lệ đó cao nhất là bao nhiêu?

Muốn nhận được ước lượng khoảng chỉ tỷ lệ thực với độ chính xác là 0,03 thì cần thăm dò thêm bao nhiêu phiếu nữa?

Với mẫu $n = 300$, ước lượng khoảng có độ chính xác là 0,0436 thì độ tin cậy của kết luận về ước lượng khoảng đó là bao nhiêu?

Bài 14. Dùng phương pháp hấp thụ nguyên tử để phân tích lượng kẽm trong tóc, một cán bộ đã phân tích 35 mẫu tóc. Kết quả thu được như sau: X là lượng kẽm trong tóc, đơn vị đo là ppm (phần triệu): Giả sử lượng kẽm trong tóc tuân theo phân bố chuẩn.

X	188	190	193	195	196	198	199	204
Số mẫu	3	4	5	10	7	3	2	1

1. Hãy ước lượng lượng kẽm trung bình (EX) chứa trong tóc.
2. Với độ tin cậy 0,95 có thể nói lượng kẽm trung bình thuộc khoảng nào? cao nhất là bao nhiêu? thấp nhất là bao nhiêu? khả năng đúng là bao nhiêu? khả năng sai là bao nhiêu?
3. Độ chính xác của ước lượng khoảng nhận được ở trên là bao nhiêu? Muốn độ chính xác là 3ppm thì cao phân tích bao nhiêu mẫu tóc?
4. Với mẫu đã cho ban đầu, ước lượng khoảng cho EX có độ chính xác là 2ppm thì độ tin cậy là bao nhiêu?

2 Chương 2: Kiểm định

Bài 1. Trung tâm hỗ trợ người tiêu dùng nhận được khá nhiều lời phàn nàn về sản phẩm bột giặt loại 4 Kg của công ty Sáng Chói. Để hỗ trợ người tiêu dùng, Trung tâm tiến hành chọn ngẫu nhiên 36 gói bột giặt của công ty để cân và thu được kết quả trung bình mẫu 3,95 Kg. Giả sử trọng lượng bột giặt sản xuất của công ty tuân theo quy luật phân phối chuẩn với độ lệch chuẩn là 0,15 Kg.

1. Trung tâm có kết luận gì khi thực hiện kiểm định giả thuyết với mức ý nghĩa 5%
2. Trung tâm có kết luận gì khi thực hiện kiểm định giả thuyết với mức ý nghĩa 2%

Bài 2. Trọng lượng (X) sản phẩm do nhà máy sản xuất ra là biến ngẫu nhiên phân phối chuẩn với độ lệch chuẩn $\sigma = 2$ (kg) và trọng lượng trung bình là 20 (kg). Nghi ngờ máy hoạt động không bình thường làm thay đổi trọng lượng trung bình của sản phẩm người ta cân thử 100 sản phẩm và thu được kết quả sau:

Trọng lượng sản phẩm	19	20	21	22	23
Số sản phẩm	10	50	20	15	5

Với mức ý nghĩa 0,05 hãy kết luận về điều nghi ngờ trên.

Bài 3. Mỳ chính được đóng gói 453 gam một gói trên máy tự động. Có thể coi trọng lượng các gói mỳ chính tuân theo quy luật chuẩn với độ lệch chuẩn 36 gam. Kiểm tra ngẫu nhiên 81 gói thấy trọng lượng trung bình là 448 gam. Với mức ý nghĩa $\alpha = 0,05$ có thể kết luận trọng lượng các gói mỳ chính có xu hướng bị đóng thiếu không?

Bài 4. Một nhà máy sản xuất bánh ngọt tuyên bố rằng mỗi chiếc bánh của họ trung bình có 88 calo. Một mẫu ngẫu nhiên với 46 chiếc bánh được kiểm tra cho thấy lượng calo trung bình trong mỗi chiếc bánh là 90 calo với độ lệch tiêu chuẩn là 4 calo. Với mức ý nghĩa 5%, kiểm định xem có phải trên thực tế mỗi chiếc bánh về trung bình chứa nhiều hơn 88 calo hơn hay không?

Bài 5. Năng suất lúa trung bình của giống lúa A được công bố là 43 tạ/ha. Một nhóm gồm 60 thửa ruộng thí nghiệm được kiểm tra cho thấy năng suất lúa trung bình của nhóm là 46,2 tạ/ha với độ lệch chuẩn 12 tạ/ha. Với mức ý nghĩa 5%, nhận định xem có phải công bố là thấp hơn so với sự thật không?

Bài 6. Trong điều kiện bình thường trọng lượng trung bình sản phẩm do nhà máy sản xuất ra là 20 kg. Nghi ngờ máy hoạt động không bình thường làm thay đổi trọng lượng trung bình của sản phẩm người ta cân thử 100 sản phẩm và thu được kết quả sau:

Trọng lượng sản phẩm	19	20	21	22	23
Số sản phẩm	10	60	20	5	5

Với mức ý nghĩa 0,05 hãy kết luận về điều nghi ngờ trên.

Bài 7. Trọng lượng của con gà lúc mới nở là đại lượng ngẫu nhiên tuân theo quy luật phân phối chuẩn. Nghi ngờ về độ đồng đều về trọng lượng gà con giảm sút người ta cân thử 12 con và tìm được $s^2 = 11,41(gam^2)$. Với mức ý nghĩa $\alpha = 0,05$, hãy kết luận về điều nghi ngờ trên biết rằng bình thường độ phân tán của trọng lượng gà con là $\sigma^2 = 10(gam^2)$.

Bài 8. Nếu máy móc hoạt động bình thường thì trọng lượng sản phẩm tuân theo luật phân phối chuẩn với độ lệch chuẩn $\sigma = 1$ (kg). Có thể nói máy móc còn hoạt động tốt hay không nếu cân thử 30 sản phẩm ta thấy độ lệch chuẩn mẫu tăng lên tới 1,1 (kg). Yêu cầu với mức ý nghĩa $\alpha = 0,01$.

Bài 9. Lô hàng đủ tiêu chuẩn xuất khẩu nếu tỷ lệ phế phẩm không vượt quá 3%. Kiểm tra ngẫu nhiên 400 sản phẩm của lô hàng thấy có 14 phế phẩm. Với mức ý nghĩa $\alpha = 0,05$ có cho phép lô hàng xuất khẩu được hay không?

Bài 10. Tỷ lệ phế phẩm do một nhà máy tự động sản xuất là 5%. Kiểm tra ngẫu nhiên 300 sản phẩm thấy có 24 phế phẩm. Nên có ý kiến cho rằng tỷ lệ phế phẩm do nhà máy sản xuất có chiều hướng tăng lên. Hãy kết luận ý kiến trên với mức ý nghĩa $\alpha = 0,05$.

Bài 11. Một tỉnh báo cáo tỉ lệ học sinh tốt nghiệp của họ là 88%. Một mẫu ngẫu nhiên 100 học sinh được chọn thì chỉ có 82 em đỗ. Với mức ý nghĩa 5% kiểm định xem báo cáo của tỉnh có cao hơn sự thật.

Bài 12. Tại thành phố M, mỗi hộ dùng không quá một điện thoại bàn và các điện thoại bàn chỉ sử dụng dịch vụ của một trong 3 công ty viễn thông A, B và C. Điều tra ngẫu nhiên 3600 hộ tại thành phố M thấy có 2500 hộ dùng điện thoại bàn, trong đó có 1300 hộ dùng điện thoại bàn sử dụng dịch vụ viễn thông của công ty A.

1. Với độ tin cậy 95% hãy ước lượng khoảng tỉ lệ hộ dùng điện thoại bàn tại thành phố M.
2. Với mức ý nghĩa 1% có thể cho rằng số điện thoại bàn sử dụng dịch vụ viễn thông của công ty A nhiều hơn hai công ty còn lại không?

Bài 13. Để so sánh trọng lượng trung bình của trẻ sơ sinh ở thành thị và nông thôn người ta cân thử 1000 trẻ ở hai khu vực và thu được số liệu:

Vùng	Số trẻ được cân	Trung bình mẫu	Độ lệch chuẩn mẫu
Nông thôn	800	3,0 kg	0,3 kg
Thành thị	200	3,2 kg	0,3 kg

Với mức ý nghĩa $\alpha = 0,05$ có thể coi trọng lượng trung bình của trẻ sơ sinh ở thành thị cao hơn ở nông thôn hay không? Giả thiết trọng lượng trẻ sơ sinh có phân phối chuẩn.

Bài 14. Người ta nghiên cứu năng suất lúa mỳ ở hai vùng chế độ canh tác khác nhau, kết quả thu được như sau:

Vùng	Số thửa ruộng canh tác	Trung bình mẫu	Phương sai mẫu
A	9	24,6 tạ/ha	0,24 (tạ/ha) ²
B	16	25,8 tạ/ha	0,16 (tạ/ha) ²

Với mức ý nghĩa $\alpha = 0,05$ hỏi có sự khác nhau đáng kể về năng suất lúa trung bình giữa hai vùng đất canh tác không?

Bài 15. Kiểm tra chất lượng của hai lô sản phẩm người ta thấy ở lô thứ nhất trong 500 sản phẩm được kiểm tra có 50 phế phẩm, ở lô thứ hai trong 400 sản phẩm được kiểm tra có 60 phế phẩm. Với mức ý nghĩa $\alpha = 0,05$ có thể xem tỉ lệ phế phẩm của hai lô hàng bằng nhau không?

Bài 16. Độ tinh khiết của một chất xúc tác rất quan trọng trong nghiên cứu hóa học. Người ta thử nghiệm hai phương pháp khác nhau: bằng phương pháp I (hữu cơ) làm 34 mẫu và bằng phương pháp II (vô cơ) làm 22 mẫu. Kết quả thu được như sau (lượng chất bần trên một đơn vị chất):

Phương pháp I

2,0 2,0 1,8 0,9 1,7 1,6 1,7 1,5 1,9 2,0 1,8 1,6 1,8 1,7 2,1 1,5

1,7 2,0 1,8 1,7 1,5 1,6 1,6 1,7 1,7 1,4 1,5 1,7 1,6 2,0 1,9 2,1

Phương pháp II

1,5 1,4 1,5 1,6 1,1 1,7 1,4 1,7 1,4 1,4 1,7 1,1 1,5 1,2 2,0 1,6

1,1 1,3 1,5 1,7 1,9 1,0

Với mức ý nghĩa $\alpha = 0,05$ có thể coi lượng chất bần trung bình của hai phương pháp trên là khác nhau được không? Giả sử rằng lượng chất bần trên một đơn vị chất có phân phối chuẩn.

Bài 17. Người ta làm một bài kiểm tra trên 15 sinh viên để tìm hiểu hiệu quả của việc giảng dạy một vấn đề theo phương pháp mới. Trước khi học phương pháp này sinh viên được cho làm một bài kiểm tra và sau khi học làm bài kiểm tra thứ 2 (điểm chấm theo thang điểm 100). Kết quả thu được như sau:

Số thứ tự	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Trước học	54	79	91	75	68	43	33	85	22	56	73	63	29	75	87
Sau học	66	85	83	88	93	40	78	91	44	82	59	81	64	83	81

Với mức ý nghĩa $\alpha = 0,05$ có thể nói rằng học theo phương pháp mới hiệu quả hơn không? Giả thiết điểm số của sinh viên có phân phối chuẩn.

Bài 18. Hai máy tiện như nhau, nhưng hoạt động trong các điều kiện thời tiết khác nhau. Sau một thời gian sản xuất người ta nghi ngờ chất lượng hoạt động của chúng khác nhau. Điều đó có đúng không nếu trong 1000 sản phẩm do máy I làm ra có 140 phế phẩm, còn trong 2000 sản phẩm do máy II làm ra có 260 phế phẩm. Hãy kết luận điều nghi ngờ trên với mức ý nghĩa $\alpha = 0,05$

Bài 19. Nghiên cứu trọng lượng của trẻ em lứa tuổi lên 10 ở thành phố và nông thôn, ta có hai mẫu đại diện sau

kg	< 35	35 – 38	38 – 41	41 – 44	44 – 47	47 – 50	50 – 53	≥ 53
TP	0	2	8	13	20	15	12	8
NT	5	10	12	15	10	3	0	0

a. Với mức ý nghĩa $\alpha = 0,05$ có thể nói trọng lượng trung bình của trẻ em lứa tuổi lên 10 ở hai vùng trên như nhau hay không? hay vùng nào cao hơn?

b. Ở lứa tuổi này ta xem trọng lượng ≥ 50 kg là diện thừa cân, có nguy cơ béo phì. Từ số liệu trên có thể kết luận: Tỷ lệ trẻ em lên 10 ở thành phố có nguy cơ béo phì cao hơn 25% hay không (với mức ý nghĩa $\alpha = 0.1$)?

Bài 20. Thống kê cân nặng của 10 cặp cha con. Kết quả thu được như sau (kg) :

Cha (X) (kg)	60	62	64	66	68	70	71	72	74	76
Con(Y) (kg)	64	63.5	65	67	69.5	71	72.5	73	74.5	76.5

Hãy ước lượng hệ số tương quan X và Y và tìm đường hồi quy tuyến tính thực nghiệm của Y theo X.