

- Tên học phần: Thống kê sinh học
- Mã học phần: SIN2122 - Số tín chỉ: 2
- Thời gian làm bài: 90 phút (không kể thời gian chép/phát đề)
- Loại đề: Không được sử dụng tài liệu

Câu 1 (2 điểm): Đánh giá sự phân bố mật độ của hai loài sao biển tại 4 trạm khảo sát ký hiệu từ ST1 - ST4, người ta thu được kết quả như sau (đơn vị: cá thể/100m²):

Loài	ST1	ST2	ST3	ST4
<i>Asterias sp1.</i>	250	45	30	80
<i>Asterias sp2.</i>	50	5	10	6

Với mức ý nghĩa $\alpha = 0,05$, anh (chị) hãy so sánh tần số xuất hiện của hai loài sao biển trên. Biết $\chi^2_{0,05, 3df} = 7,815$.

Câu 2 (3 điểm): Khi nghiên cứu hình thái một loài thực vật, người ta tiến hành đo chiều dài lá của loài này và nhận thấy có sự khác nhau về thông số này khi cây sinh trưởng ở 3 vùng khác nhau (ký hiệu là vùng A, B và C). Các số đo thu thập được như sau (đơn vị: cm):

Vùng A	12,1	11,0	13,2	11,6	12,4	11,9	11,9	11,2	11,3	12,8	13,0	14,1	
Vùng B	11,1	12,1	12,0	12,8	11,0	10,9	12,1	12,5	12,2	11,1	10,8	12,0	11,7
Vùng C	14,1	13,2	14,5	12,3	13,7	14,0	13,2	15,0	13,2	14,2			

Đặt giả thuyết $H_0: M_1 = M_2 = M_3$; H_A : có ít nhất một cặp giá trị trung bình khác nhau; mức ý nghĩa được chọn $\alpha=0,05$, sử dụng phân tích phương sai một nhân tố (ANOVA: Single Factor) cho kết quả sau:

SUMMARY

Groups	Count	Sum	Average	Variance
Vùng A	12	146,5	12,2	0,86
Vùng B	13	152,3	11,7	0,44
Vùng C	10	137,4	13,7	0,61

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	24,35	2	12,2	19,26	0,000003	3,29
Within Groups	20,23	32	0,6			
Total	44,58	34				

- Viết nhận xét kết quả phân tích ANOVA.
- Sử dụng Scheffe's test để so sánh sự sai khác về chiều dài lá cây ở các vùng.

Câu 3 (3 điểm): Để nghiên cứu sự khác nhau giữa trọng lượng tôm sú nuôi tại hai ao khác nhau, người ta bắt ngẫu nhiên mỗi ao 12 con, đem cân và thu được kết quả như sau (đơn vị: g)

Ao 1	33	31	33	34	32	28	36	32	34	29	34	35
Ao 2	35	32	28	37	39	35	36	33	37	38	38	34

a) Anh chị hãy đặt giả thiết để kiểm định sự khác biệt về phương sai mẫu.

Với mức ý nghĩa $\alpha=0,05$, sử dụng F-test thu được:

	Ao 2	Ao 1
Mean	35,2	32,6
Variance	9,61	5,54
Observations	12	12
Df	11	11
F	1,735	
P(F<=f) one-tail	0,187	
F Critical one-tail	2,818	

b) Viết giả thiết kiểm định T-test.

Sử dụng T-test phù hợp, với mức ý nghĩa $\alpha=0,05$ thu được kết quả:

	Ao 1	Ao 2
1 Mean	32,6	35,2
2 Variance	5,54	9,61
3 Observations	12	12
4 Pooled Variance	7,6	
5 Hypothesized Mean Difference	0	
6 Df	22	
7 t Stat	-2,300	
P(T<=t) one-tail	0,016	
t Critical one-tail	1,717	
P(T<=t) two-tail	0,031	
t Critical two-tail	2,074	

Anh (chị) hãy viết nhận xét kết quả phân tích.

Câu 4 (2 điểm): Phân tích mối tương quan giữa thời gian (tính theo giờ) và hoạt độ (mol/giờ) của một chế phẩm sinh học được số liệu sau:

Thời gian	0	1	2	3	4	5	6	7	8
Hoạt độ	28,40	25,5	23,48	19,59	15,79	13,28	12,65	10,89	9,54

tương quan tuyến tính được xác định bằng lệnh correlation trong excel là $r = -0,983$
 Kết quả phân tích hồi quy tuyến tính bằng MS Excel thu được:

Hệ số

14,165

2,6

SUMMARY OUTPUT

Regression Statistics

Multiple R	0,983
R Square	0,966
Adjusted R Square	0,961
Standard Error	1,343
Observations	9

ANOVA

	df	SS	MS	F	Significance F
Regression	1	361,3	361,3	200,3	2,09E-06
Residual	7	12,6	1,8		
Total	8	374,0			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%
Intercept	27,50	0,83	33,31	5,7E-09	25,54
Thời gian	-2,45	0,17	-14,15	2,09E-06	-2,86

Anh chị hãy

a) Xác định sự tồn tại của hệ số tương quan tuyến tính giữa X và Y, viết nhận xét về mối tương quan này. Cho biết $t_{0,05, 7df} = 2,365$.

b) Viết phương trình hồi quy tuyến tính và tìm hoạt độ chế phẩm ở thời gian 5,5 giờ.

(Ghi chú: Cán bộ coi thi không được giải thích gì thêm)