ΤΜΗΜΑ ΓΕΩΠΟΝΙΑΣ-ΑΓΡΟΤΕΧΝΟΛΟΓΙΑΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

**Λύσεις 2ης ΕΡΓΑΣΙΑΣ ΒΙΟΜΕΤΡΙΑ-ΓΕΩΡΓΙΚΟΣ ΠΕΙΡΑΜΑΤΙΣΜΟΣ**

***Άσκηση 2***

Δίνεται το μήκος ίνας του βαμβακιού 10 μητρικών φυτών και του μέσου όρου των απογόνων τους. Δεδομένου ότι το μήκος της ίνας κληρονομείται, είναι λογικό να θεωρήσουμε ότι το μήκος της ίνας των απογόνων καθορίζεται από το μήκος της ίνας των γονέων, ότι δηλαδή το μήκος της ίνας των γονέων είναι το αίτιο και το μήκος της ίνας των απογόνων είναι το αποτέλεσμα.

|  |  |
| --- | --- |
| **Μητρικό φυτό (X)** | **Απόγονοι (Y)** |
| 26 | 24,3 |
| 28,8 | 24,5 |
| 25,2 | 23,4 |
| 23,4 | 21,6 |
| 26,6 | 22,5 |
| 25,4 | 23,6 |
| 23,4 | 22,8 |
| 27,6 | 24,4 |
| 24,4 | 22,6 |
| 24 | 25,4 |

Υπολογίσετε την εξίσωση που δίνει το μήκος της ίνας των απογόνων συναρτήσει του μήκους της ίνας του μητρικού φυτού, ελέγξτε την σημαντικότητα του συντελεστού συνμεταβολής και υπολογίστε τα όρια εμπιστοσύνης του. Κάνετε ένα Scaterplot με την γραμμή συνμεταβολής.

***Λύση :***

Για να βρεθεί η σχέση ανάμεσα σε ενα μητρικό φυτό βαμβακιού (Χ) των απογόνων του (Υ) εφαρμόστηκε το γραμμικό μοντέλο της συνμεταβολής.



Το γραμμικό μοντέλο συνμεταβολής είναι : y=16,42+ 0,28\*x και ο συντελεστής προσδιορισμού (R square) R2= 0,188 o οποίος εκφράζει το ποσοστό της παραλλακτικότητας που ερμηνεύει το μοντέλο το οποίο σε αυτή τη περίπτωση είναι 18,8 %.

Στη συνέχεια, ελέγχουμε αν ισχύουν οι παραδοχές (assumptions) του γραμμικού μοντέλου συνμεταβολής:

* Δεν υπάρχουν ακραίες τιμές αφού σύμφωνα με το παρακάτω πίνακα οι ελάχιστες και οι μέγιστες τιμές των υπολειμμάτων του σφάλματος (πίνακας residual statistics) είναι minimum -1,207 και maximum 2,087 αντίστοιχα δηλαδή μικρότερες από το ±3,29 τυπικές αποκλίσεις.

|  |
| --- |
| **Residuals Statisticsa** |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | 22,9309 | 24,4343 | 23,5100 | ,50095 | 10 |
| Residual | -1,33089 | 2,30206 | ,00000 | 1,03979 | 10 |
| Std. Predicted Value | -1,156 | 1,845 | ,000 | 1,000 | 10 |
| Std. Residual | -1,207 | 2,087 | ,000 | ,943 | 10 |
| a. Dependent Variable: Απόγονοι (Y) |
|  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| * Έλεγχος ανεξαρτησίας υπολειμμάτων σφάλματος. Τα υπολείμματα του σφάλματος ήταν ανεξάρτητα μεταξύ τους καθώς ο δείκτης Durbin-Watson είναι 1,401 δηλαδή η τιμή βρίσκεται ανάμεσα στο 1 και το 3.

|  |
| --- |
| **Model Summaryb** |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | ,434a | ,188 | ,087 | 1,10286 | 1,401 |
| a. Predictors: (Constant), Μητρικό Φυτό (Χ) |
| b. Dependent Variable: Απόγονοι (Y) |

 |

* Σύμφωνα με το ιστόγραμμα (όπου υπάρχει καμπύλη σχήματος καμπανας) και το pplot (όπου οι τυποποιημένες δεν απεχουν πολύ από την γραμμη κλισεως 45°), παρατηρούμε πως η κατανομή των σφαλμάτων είναι κανονική. 



* Η συνμεταβολή είναι στατιστικώς μη σημαντική αφού F: Sig = 0,210>0,05, όπως εμφανίζεται στον πίνακα ANOVA.

Ακόμη το όριο εμπιστοσύνης (95%) του b (-0,193 έως 0,750) δηλώνουν ότι αφού συμπεριλαμβάνεται στα όρια η τιμή 0 δεν συνμεταβάλλονται ταυτόχρονα σε όλο το εύρος των τιμών. Επίσης επειδή ο συντελεστής προσδιορισμού είναι πολύ χαμηλός (R2=0,188) δηλαδή το ποσοστό παραλλακτικότητας που ερμηνεύεται είναι 18,8 % ένα πολύ μικρό ποσοστό της παραλλακτηκότητας ερμηνεύει το γραμμικό μοντέλο συνμεταβολής και δεν δεχόμαστε τη σχέση ως σημαντική.

|  |
| --- |
| **ANOVAa** |
| Model | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 2,259 | 1 | 2,259 | 1,857 | ,210b |
| Residual | 9,730 | 8 | 1,216 |  |  |
| Total | 11,989 | 9 |  |  |  |
| a. Dependent Variable: Απόγονοι (Y) |
| b. Predictors: (Constant), Μητρικό Φυτό (Χ) |

|  |
| --- |
| **Coefficientsa** |
| Model | Unstandardized Coefficients | Standardized Coefficients | t | Sig. | 95,0% Confidence Interval for B |
| B | Std. Error | Beta | Lower Bound | Upper Bound |
| 1 | (Constant) | 16,416 | 5,218 |  | 3,146 | ,014 | 4,384 | 28,448 |
| Μητρικό Φυτό (Χ) | ,278 | ,204 | ,434 | 1,363 | ,210 | -,193 | ,750 |
|  |  |  |  |  |  |  |  |  |
| a. Dependent Variable: Απόγονοι (Y) |

***Άσκηση 3***

Η απόδοση της σουλτανίνας σε σταφίδα εξαρτάται από τον βαθμό της ωριμότητας της. Τα παρακάτω δεδομένα είναι ο βαθμός ωριμότητας των σταφυλιών (Χ) και η επί τοις εκατό απόδοση σε σταφίδα (Υ).

|  |  |
| --- | --- |
| **Ωριμότητα (Χ)**  | **Απόδοση % (Υ)**  |
| 9,8 | 20,6 |
| 10,8 | 22,6 |
| 11,6 | 24,3 |
| 12,2 | 26,1 |
| 12,8 | 27,2 |
| 13,4 | 28,4 |
| 13,7 | 29,1 |
| 14,9 | 31,8 |
| 15,4 | 33 |
| 16,4 | 35,4 |

Υπολογίσετε την εξίσωση που δίνει την σχέση μεταξύ ωριμότητας και απόδοσης, ελέγξετε την σημαντικότητα του b, και υπολογίστε τα όρια εμπιστοσύνης του. Κάνετε ένα Scaterplot με την γραμμή συνμεταβολής.

***Λύση:***

Για να μελετηθεί η σχέση μεταξύ ωριμότητας σταφυλιών (Χ) ποικιλίας σουλτανίνας και της % απόδοσης της (Υ)εφαρμόστηκε το γραμμικό μοντέλο της συνμεταβολής.



Όπως εμφανίζεται στο Διγράφημα Διασποράς τιμών η εξίσωση του μοντέλου είναι: y=1,52+2,24\*x και ο συντελεστής προσδιορισμού είναι R2= 0,999 ο οποίος εκφράζει το ποσοστό της παραλλακτικότητας που ερμηνεύει το μοντέλο και είναι 99,9 %.

Ελέγχουμε αν ισχύουν οι παραδοχές (assumptions) που διέπουν την εφαρμογή του μοντέλου στα δεδομένα:

* Ελέγχουμε για outliers (ακραίες τιμές) στον πίνακα **Residuals Statistics**. Σύμφωνα με τον παρακάτω πίνακα παρατηρούμε πως οι ελάχιστες και οι μέγιστες τιμές των residual statistics είναι -1,207 και 1,726 αντίστοιχα δηλαδή μικρότερες από το ±3,29, κάτι που μας οδηγεί στο συμπέρασμα πως δεν υπάρχουν ακραίες τιμές.

|  |
| --- |
| **Residuals Statisticsa** |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | 20,4519 | 35,2481 | 27,8500 | 4,67871 | 10 |
| Residual | -,18724 | ,26765 | ,00000 | ,14621 | 10 |
| Std. Predicted Value | -1,581 | 1,581 | ,000 | 1,000 | 10 |
| Std. Residual | -1,207 | 1,726 | ,000 | ,943 | 10 |
| a. Dependent Variable: Απόδοση%(Y) |
| * Έλεγχος ανεξαρτησίας υπολειμμάτων σφάλματος. Οι μετρήσεις ήταν ανεξάρτητες μεταξύ τους αφού ο δείκτης Durbin-Watson είναι 2,014 δηλαδή η τιμή βρίσκεται ανάμεσα στο 1 και το 3.

|  |
| --- |
| **Model Summaryb** |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | 1,000a | ,999 | ,999 | ,15508 | 2,014 |
| a. Predictors: (Constant), Ωριμότητα(X) |
| b. Dependent Variable: Απόδοση%(Y) |

 |

* Η κατανομή των σφαλμάτων είναι κανονική καθώς παρατηρούμε πως στο ιστόγραμμα (η καμπύλη συχνοτήτων υπολειμμάτων σφάλματος έχει κατανομή σε σχήμα καμπάνας) και στο pplot (όπου οι κουκίδες δεν απέχουν πολύ από την γραμμή κλίσης 45°).





Κατόπιν ελέγχουμε τον πίνακα της ANOVA ως προς τη σημαντικότητα του F. Η συνμεταβολή είναι στατιστικώς σημαντική αφού F: Sig = 0,000<0,05. Επίσης το όριο εμπιστοσύνης (95%) του b (2,185 εως 2,299) δηλώνουν ότι συνμεταβάλλονται ταυτόχρονα για κάθε τιμή του (X), εφόσον δεν συμπεριλαμβάνεται στα όρια η τιμή 0. Ακόμα το R2=0,999 δηλαδή το ποσοστό παραλλακτικότητας που ερμηνεύει το μοντέλο είναι 99,9% πράγμα που μας οδηγεί στο συμπέρασμα ότι είναι ισχυρή η σχέση μεταξύ ωριμότητας σταφυλιών (Χ) σουλτανίνας και της % απόδοσης της (Υ) (y=1,52+2,24\*x) και η ωριμότητα μπορεί να προβλέψει την απόδοση των σταφυλιών σουλτανίνας. Συμπερασματικά για κάθε 1 μονάδα αύξηση στο βαθμό ωριμότητας η % απόδοση αυξάνεται κατά 2,24 μονάδες.

|  |
| --- |
| **ANOVAa** |
| Model | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 197,013 | 1 | 197,013 | 8192,314 | ,000b |
| Residual | ,192 | 8 | ,024 |  |  |
| Total | 197,205 | 9 |  |  |  |
| a. Dependent Variable: Απόδοση%(Y) |
| b. Predictors: (Constant), Ωριμότητα(X) |

|  |
| --- |
| **Coefficientsa** |
| Model | Unstandardized Coefficients | Standardized Coefficients | t | Sig. | 95,0% Confidence Interval for B |
| B | Std. Error | Beta | Lower Bound | Upper Bound |
| 1 | (Constant) | -1,518 | ,328 |  | -4,626 | ,002 | -2,275 | -,761 |
| Ωριμότητα(X) | 2,242 | ,025 | 1,000 | 90,511 | ,000 | 2,185 | 2,299 |
| a. Dependent Variable: Απόδοση%(Y) |
| ***Άσκηση 5***Τα παρακάτω δεδομένα είναι οι αποδόσεις σε χόρτο (kg/φυτό) 12 μεμονωμένων φυτών μηδικής κατά την πρώτη και τη δεύτερη κοπή.

|  |  |
| --- | --- |
| **Πρώτη κοπή**  | **Δεύτερη κοπή**  |
| 0,3 | 1,1 |
| 0,46 | 0,99 |
| 0,31 | 1,31 |
| 0,47 | 0,75 |
| 0,41 | 1,22 |
| 0,07 | 0,4 |
| 0,58 | 1,66 |
| 0,52 | 1,58 |
| 0,4 | 0,84 |
| 0,33 | 1,18 |
| 0,26 | 1,1 |
| 0,46 | 1,25 |

Οι μεταβλητές είναι ανεξάρτητες μεταξύ τους. Υπολογίστε το συντελεστή συσχετίσεως μεταξύ της πρώτης και δεύτερης κοπής και ελέγξετε τη σημαντικότητα του.***Λύση :***Για να υπολογίσουμε τον συντελεστή συσχετίσεως μεταξύ της πρώτης κοπής των μηδικών φυτών και της δεύτερης θα πρέπει να ελέγξουμε πρώτα εάν οι μεταβλητές κατανέμονται κανονικά και στη συνέχεια να υπολογίσουμε την σημαντικότητα του συντελεστή συσχέτισης r (Pearson). |

Σύμφωνα με τα δεδομένα οι Skewness της ***πρώτης κοπής*** έχει τιμές -0,871 έως 0,637 δηλαδή εντός των ορίων ±1 και Kyrtossis 1,225 έως 1,232 δηλαδή εντός των ορίων ±2. Ακόμη σύμφωνα με τα δεδομένα οι Skewness της ***δεύτερης κοπής*** έχει τιμές -0,428 έως 0,637 και Kyrtossis 0,578 έως 1,232 δηλαδή εντός των ορίων. Συμπερασματικά τα δεδομένα κατανέμονται κανονικά.

|  |
| --- |
| **Descriptives** |
|  | Statistic | Std. Error |
| Πρώτη Κοπή | Mean | ,3808 | ,03951 |
| 95% Confidence Interval for Mean | Lower Bound | ,2939 |  |
| Upper Bound | ,4678 |  |
| 5% Trimmed Mean | ,3870 |  |
| Median | ,4050 |  |
| Variance | ,019 |  |
| Std. Deviation | ,13688 |  |
| Minimum | ,07 |  |
| Maximum | ,58 |  |
| Range | ,51 |  |
| Interquartile Range | ,16 |  |
| Skewness | -,871 | ,637 |
| Kurtosis | 1,225 | 1,232 |
| Δεύτερη Κοπή | Mean | 1,1150 | ,10003 |
| 95% Confidence Interval for Mean | Lower Bound | ,8948 |  |
| Upper Bound | 1,3352 |  |
| 5% Trimmed Mean | 1,1244 |  |
| Median | 1,1400 |  |
| Variance | ,120 |  |
| Std. Deviation | ,34653 |  |
| Minimum | ,40 |  |
| Maximum | 1,66 |  |
| Range | 1,26 |  |
| Interquartile Range | ,42 |  |
| Skewness | -,428 | ,637 |
| Kurtosis | ,578 | 1,232 |

H συσχέτιση Pearson μεταξύ της απορρόφησης από το θρεπτικό διάλυμα και το χώμα είναι r=0,66 στατιστικώς σημαντική (Πιθανότητα σφάλματος =0,018 δλδ 1,8% επειδή 0,018<0,05 δλδ 5%).

Συμπερασματικά η συσχέτιση ανάμεσα στην πρώτη κοπή και την δεύτερη είναι r=0,66 και στατιστικώς σημαντική.

|  |
| --- |
| **Correlations** |
|  | Πρώτη Κοπή | Δεύτερη Κοπή |
| Πρώτη Κοπή | Pearson Correlation | 1 | ,666\* |
| Sig. (2-tailed) |  | ,018 |
| N | 12 | 12 |
| Δεύτερη Κοπή | Pearson Correlation | ,666\* | 1 |
| Sig. (2-tailed) | ,018 |  |
| N | 12 | 12 |
| \*. Correlation is significant at the 0.05 level (2-tailed). |

***Άσκηση 6***

Στον παρακάτω πίνακα είναι οι περιεκτικότητες σε Ca του γάλακτος δυο περιοχών Α και Β.

|  |  |
| --- | --- |
| **Περιοχή Α** | **Περιοχή Β** |
| 124,8 | 125,9 |
| 116,2 | 122,7 |
| 116,5 | 119,6 |
| 125,8 | 119,6 |
| 120,7 | 122,7 |
| 124,3 | 123,9 |
| 126,2 | 128,4 |
| 128,9 | 128,4 |
| 123,6 | 130,0 |
| 124,4 | 129,0 |
| 121,0 | 118,5 |
| 120,3 | 122,2 |

Οι μεταβλητές είναι ανεξάρτητες μεταξύ τους. Υπολογίστε το συντελεστή συσχετίσεως μεταξύ της πρώτης και δεύτερης περιοχής και ελέγξετε τη σημαντικότητα του.

***Λύση:***

Αρχικά πρέπει να ελέγξουμε εάν οι μεταβλητές κατανέμονται κανονικά και κατόπιν να υπολογίσουμε to συντελεστή συσχέτισης r (Pearson) και την σημαντικότητα του.

Σύμφωνα με τα δεδομένα οι Skewness της περιοχής Α έχει τιμές -0,420 έως 0,637 δηλαδή εντός των ορίων ±1 και Kyrtossis -0,433 έως 1,232 δηλαδή εντός των ορίων ±2. Ακόμη σύμφωνα με τα δεδομένα οι Skewness της περιοχής Β έχει τιμές 0,070 έως 0,637 και Kyrtossis -1,475 έως 1,232 δηλαδή εντός των ορίων. Συμπερασματικά τα δεδομένα κατανέμονται κανονικά.

|  |
| --- |
| **Descriptives** |
|  | Statistic | Std. Error |
| Περιοχή Α | Mean | 122,7250 | 1,11567 |
| 95% Confidence Interval for Mean | Lower Bound | 120,2694 |  |
| Upper Bound | 125,1806 |  |
| 5% Trimmed Mean | 122,7444 |  |
| Median | 123,9500 |  |
| Variance | 14,937 |  |
| Std. Deviation | 3,86479 |  |
| Minimum | 116,20 |  |
| Maximum | 128,90 |  |
| Range | 12,70 |  |
| Interquartile Range | 5,15 |  |
| Skewness | -,420 | ,637 |
| Kurtosis | -,433 | 1,232 |
| Περιοχή Β | Mean | 124,2417 | 1,16199 |
| 95% Confidence Interval for Mean | Lower Bound | 121,6841 |  |
| Upper Bound | 126,7992 |  |
| 5% Trimmed Mean | 124,2407 |  |
| Median | 123,3000 |  |
| Variance | 16,203 |  |
| Std. Deviation | 4,02525 |  |
| Minimum | 118,50 |  |
| Maximum | 130,00 |  |
| Range | 11,50 |  |
| Interquartile Range | 8,15 |  |
| Skewness | ,070 | ,637 |
| Kurtosis | -1,475 | 1,232 |

 Κάνοντας την ανάλυση η συσχέτιση μεταξύ της απορρόφησης από το θρεπτικό διάλυμα και το χώμα είναι r=0,577 στατιστικώς σημαντική Πιθανότητα=0,05 (δλδ σφάλμα αποδεκτό 5%).

Συμπερασματικά η συσχέτιση ανάμεσα στις περιεκτικότητες σε Ca του γάλακτος δυο περιοχών Α και Β είναι μέση r=0,57 και στατιστικώς σημαντική.

|  |
| --- |
| **Correlations** |
|  | Περιοχή Α | Περιοχή Β |
| Περιοχή Α | Pearson Correlation | 1 | ,577\* |
| Sig. (2-tailed) |  | ,050 |
| N | 12 | 12 |
| Περιοχή Β | Pearson Correlation | ,577\* | 1 |
| Sig. (2-tailed) | ,050 |  |
| N | 12 | 12 |
| \*. Correlation is significant at the 0.05 level (2-tailed). |