**Θεμα 1: Εκλογικές Περιφέρειες: Σε μια πόλη, 3 εκλογικές περιφέρειες έχουν αντίστοιχα 30%, 45% και 25% της συνολικής ψηφοφορίας. Ένας υποψήφιος κερδίζει με πιθανότητα 60%, 50% και 40% σε κάθε περιφέρεια αντίστοιχα. Ποια είναι η πιθανότητα ο υποψήφιος να κερδίσει τις εκλογές αν θεωρηθεί ότι κερδίζει εάν λάβει την πλειοψηφία της συνολικής ψήφου;**

**Λύση**

Για να υπολογίσουμε την πιθανότητα νίκης του υποψήφιου, βάσει των ποσοστών ψηφοφορίας και των πιθανοτήτων νίκης ανά περιφέρεια, χρησιμοποιούμε τον ακόλουθο τύπο:

P(win)=w1⋅p1+w2⋅p2+w3⋅p3

Στη δεδομένη περίπτωση, έχουμε:

* w1=0.30, p1=0.60 ​= 0.60 (30% των ψήφων, 60% πιθανότητα νίκης)
* w2=0.45, p2=0.50=0.50 (45% των ψήφων, 50% πιθανότητα νίκης)
* w3=0.25, p3=0.40= (25% των ψήφων, 40% πιθανότητα νίκης)

Άρα, P(win)=0.30⋅0.60+0.45⋅0.50+0.25⋅0.40=0.505

**Θεμα 2: Μια εταιρεία καταγράφει τα εβδομαδιαία έξοδα για διαφημιστικές καμπάνιες σε ευρώ κατά τη διάρκεια των τελευταίων οκτώ εβδομάδων: 500, 520, 480, 510, 530, 495, 515, 525. Υπολογίστε τον μέσο όρο, τη διάμεσο, την τυπική απόκλιση και το εύρος των εξόδων. Σχολιάστε τα αποτελέσματα και τι αυτά δείχνουν για τη σταθερότητα της επένδυσης στη διαφήμιση.**

**Λύση**

Οι διαφημιστικές δαπάνες από το παράδειγμα έχουν ως εξής: [500,520,480,510,530,495,515,525]

* **Μέσος Όρος**=(500+520+480+510+530+495+515+525)/8=509.375
* **Διάμεσος**: Οργανώνουμε τις τιμές σε αύξουσα σειρά [480, 495, 500, 510, 515, 520, 525, 530] και επειδή ο αριθμός των τιμών είναι άρτιος, ο διάμεσος είναι ο μέσος όρος των δύο μεσαίων τιμών (510+515)/2=512.5.
* **Τυπική Απόκλιση**: $s=\sqrt{\frac{\sum\_{1}^{8}\left(x\_{i}-μ\right)^{2}}{n-1}}$
* Η τυπική απόκλιση για τα δεδομένα των διαφημιστικών εξόδων είναι 16.78 ευρώ. Αυτό σημαίνει ότι οι διαφημιστικές δαπάνες διακυμαίνονται μέσω της τυπικής απόκλισης κατά περίπου 16.78 ευρώ γύρω από τον μέσο όρο τους, που είναι 509.38 ευρώ.
* **Εύρος**: Range= max(expenses) − min(expenses)=530−480=50

Ο διάμεσος βρίσκεται πολύ κοντά στον μέσο όρο, υποδηλώνοντας μια σχετικά ομοιόμορφη διακύμανση των δαπανών, χωρίς σημαντικές ακραίες τιμές ή ασυμμετρία στην κατανομή των δεδομένων.

Η τυπική απόκλιση είναι σχετικά χαμηλή σε σύγκριση με τον μέσο όρο, υποδεικνύοντας ότι οι δαπάνες δεν διακυμαίνονται σημαντικά από εβδομάδα σε εβδομάδα. Αυτό ενισχύει την έννοια της σταθερότητας στη διαφημιστική δαπάνη.

Το εύρος δείχνει τη μέγιστη διαφορά μεταξύ των υψηλότερων και χαμηλότερων εβδομαδιαίων δαπανών. Ένα εύρος των 50 ευρώ είναι σχετικά μικρό, δεδομένης της συνολικής εύρους των δαπανών, κάτι που υποδεικνύει επίσης σταθερότητα.

**Θεμα 3: Σε μια ορεινή περιοχή, η μέση ετήσια παραγωγή μηλιών ανά στρεμμα μίας συγκεκριμένου ποικιλίας είναι 50 κιλά με τυπική απόκλιση 5 κιλών. Ένας αγρότης εφάρμοσε ένα νέο φυσικό ενισχυτικό ανάπτυξης στην καλλιέργειά του και παρατήρησε ότι η μέση παραγωγή σε 120 τυχαία επιλεγμένα δεκάρια ανήλθε στα 52 κιλά. Να ελέγξετε με επίπεδο σημαντικότητας 5% αν η συγκεκριμένη αύξηση είναι στατιστικά σημαντική, δηλαδή αν το φυσικό ενισχυτικό προκαλεί πραγματικά αύξηση της μέσης παραγωγής (θεωρείστε ότι η διασπορά της παραγωγής έχει παραμείνει η ίδια). Να βρείτε το αντίστοιχο p-value της υπόθεσης.**

**H0​ (Μηδενική Υπόθεση)**: Η H0 υποθέτει ότι η μέση παραγωγή της ποικιλίας μήλων δεν έχει αλλάξει με τη χρήση του φυσικού ενισχυτικού. Δηλαδή, η μέση παραγωγή παραμένει στα 50 κιλά ανά στρέμμα. H0: μ=50

**H1​ (Εναλλακτική Υπόθεση)**: H H1​ υποθέτει ότι η μέση παραγωγή της ποικιλίας μήλων έχει αυξηθεί με τη χρήση του φυσικού ενισχυτικού. Δηλαδή, η μέση παραγωγή είναι μεγαλύτερη από 50 κιλά ανά στρέμμα. H1: μ>50

t statistic=4.38, και p value=0.000255 για 119 βαθμούς ελευθερίας

Το p-value είναι πολύ μικρότερο από το επίπεδο σημαντικότητας 0.05 (5%), που υποδηλώνει ότι η αύξηση της μέσης παραγωγής στα 52 κιλά είναι στατιστικά σημαντική. Αυτό σημαίνει ότι το νέο φυσικό ενισχυτικό ανάπτυξης φαίνεται να προκαλεί πραγματική αύξηση στην παραγωγή μήλων στη συγκεκριμένη ποικιλία.

Επομένως, μπορείτε με σιγουριά να καταλήξετε ότι η εφαρμογή του ενισχυτικού έχει επιδράσει θετικά στην παραγωγή μήλων, και η διαφορά που παρατηρήθηκε δεν οφείλεται σε τυχαίες διακυμάνσεις.

**Θέμα 4: Συνδυαστική Ερώτηση Κριτικής Σκέψης: Πώς μπορεί να αναλυθεί η επίδραση της εκπαίδευσης στην οικονομική ανάπτυξη και τις κοινωνικές διακρίσεις με βάση το φύλο, τη φυλή, ή την κοινωνική κατάσταση σε διάφορες χώρες, χρησιμοποιώντας στατιστικές μεθοδολογίες;**

Απάντηση: Για να αναλύσουμε την επίδραση της εκπαίδευσης στην οικονομική ανάπτυξη και στις κοινωνικές διακρίσεις με βάση φύλο, φυλή ή κοινωνική κατάσταση σε διάφορες χώρες, μπορούμε να εφαρμόσουμε μια σειρά από ποσοτικές και ποιοτικές μεθοδολογίες:

1. **Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression):** Χρησιμοποιούμε αυτή τη μέθοδο για να αναλύσουμε τη σχέση μεταξύ της εκπαίδευσης (π.χ., εκπαιδευτικές δαπάνες, ποσοστό αποφοίτησης) και οικονομικών δεικτών (π.χ., ΑΕΠ ανά κεφαλή). Επίσης, εξετάζουμε πώς διακρίσεις με βάση το φύλο, τη φυλή ή την κοινωνική κατάσταση επηρεάζουν την εκπαιδευτική επιτυχία ή την πρόσβαση στην εκπαίδευση.
2. **Logistic Regression:** Αναλύουμε κατηγορικά αποτελέσματα όπως την πιθανότητα πρόσβασης σε υψηλότερη εκπαίδευση ή στην αγορά εργασίας, βασιζόμενοι στις εκπαιδευτικές επιδόσεις και τις κοινωνικο-οικονομικές μεταβλητές.
3. **Ανάλυση Δεδομένων Πάνελ (Panel Data Analysis):** Χρησιμοποιείται για την εξέταση δεδομένων από πολλαπλές χρονικές περιόδους και διαφορετικές χώρες ή περιοχές, επιτρέποντας την ανάλυση της μακροχρόνιας επίδρασης των εκπαιδευτικών πολιτικών και των κοινωνικών διακρίσεων.
4. **Ποιοτική Ανάλυση Δεδομένων (Qualitative Data Analysis):** Εξετάζουμε τις προσωπικές εμπειρίες και αντιλήψεις των ατόμων σχετικά με την εκπαίδευση και τις κοινωνικές διακρίσεις, αναλύοντας δεδομένα από συνεντεύξεις, ομαδικές συζητήσεις και άλλα ποιοτικά δεδομένα.