Συντελεστής Προσδιορισμού:

Γνωρίζω ότι $SS_{tot} = SS_{reg} + SS_{res}$.

Ο συντελεστής προσδιορισμού R^2 ορίζεται από τη σχέση $R^2 = \dfrac{SS_{reg}}{SS_{tot}}$ και έχει τις εξής ιδιότητες:

(α) $0 \le R^2 \le 1$

(β) R^2 = (μεταβλητότητα που οφείλεται στην παλινδρόμηση) / (ολική μεταβλητότητα)
= ποσοστό μεταβλητότητας που οφείλεται στην παλινδρόμηση

(γ) Μικρές τιμές του R^2 παρέχουν ένδειξη "κακής" προσαρμογής.
Μεγάλες τιμές του R^2 δεν σημαίνουν κατ' ανάγκη "καλή" προσαρμογή.

Η τιμή του R^2 μπορεί να αυξηθεί εισάγοντας επιπλέον παραμέτρους στο μοντέλο.

πχ: Έστω $R_1^{\,2} = \dfrac{SS_{reg}^{\ (1)}}{SS_{tot}}$ που αντιστοιχεί στο μοντέλο $\mathrm{Y}_i = \beta_0 + \beta_1 \cdot \mathrm{X}_i + \varepsilon_i$ και $R_2^{\,2} = \dfrac{SS_{reg}^{\ (2)}}{SS_{tot}}$ που αντιστοιχεί στο μοντέλο $\mathrm{Y}_i = \beta_0 + \beta_1 \cdot \mathrm{X}_i + \beta_2 \cdot \mathrm{X}_i^{\,2} + \varepsilon_i$.

Έχω ότι $SS_{tot} = \sum_{i=1}^{n} \left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2$ και

$$SS_{reg}^{\ (1)} = SS_{tot} - SS_{res}^{\ (1)} = SS_{tot} - \min_{\beta_0,\beta_1} \sum_{i=1}^{n} \left(\mathrm{Y}_i - \beta_0 - \beta_1 \cdot \mathrm{X}_i\right)^2$$

$$SS_{reg}^{\ (2)} = SS_{tot} - SS_{res}^{\ (2)} = SS_{tot} - \min_{\beta_0,\beta_1,\beta_2} \sum_{i=1}^{n} \left(\mathrm{Y}_i - \beta_0 - \beta_1 \cdot \mathrm{X}_i - \beta_2 \cdot \mathrm{X}_i^{\,2}\right)^2$$

Αλλά, σταθεροποιώντας το $\beta_2 = 0$, έχουμε ότι

$$\min_{\beta_0,\beta_1} \sum_{i=1}^{n} \left(\mathrm{Y}_i - \beta_0 - \beta_1 \cdot \mathrm{X}_i\right)^2 \ge \min_{\beta_0,\beta_1,\beta_2} \sum_{i=1}^{n} \left(\mathrm{Y}_i - \beta_0 - \beta_1 \cdot \mathrm{X}_i - \beta_2 \cdot \mathrm{X}_i^{\,2}\right)^2$$

Άρα $SS_{reg}^{\ (2)} \ge SS_{reg}^{\ (1)} \Rightarrow R_2^{\,2} \ge R_1^{\,2}$.

Άλλος τρόπος για αύξηση της τιμής του R^2 είναι με μεγάλη διασπορά των X_i.

Ξέρουμε ότι $R^2 = \dfrac{SS_{reg}}{SS_{tot}} = \dfrac{\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i - \overline{\mathrm{Y}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2} = \dfrac{\hat{\beta}_1^{\,2} \cdot \sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2}$.

Άρα μεγάλες τιμές του R^2 είναι πιθανό να οφείλονται σε μεγάλες τιμές του $\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2$.

Γενικά, για να εξετάσω γραμμική σχέση μεταξύ X και Y, μπορώ να θεωρήσω τον συντελεστή συσχέτισης $r = \dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2}} = r(\mathrm{X},\mathrm{Y})$, που είναι εκτιμήτρια του $\rho = \dfrac{Cov(\mathrm{X},\mathrm{Y})}{\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}}$.

Για να δω αν υπάρχει γραμμική σχέση, θεωρώ τον έλεγχο $\mathrm{H}_0 : \rho = 0$ προς $\mathrm{H}_1 : \rho \neq 0$.

Όταν ισχύει η κανονικότητα, ο έλεγχος πηλίκου πιθανοφάνειας είναι: $t = \dfrac{r}{\sqrt{\dfrac{(1-r)^2}{n-2}}} \sim t_{n-2}$, όταν ισχύει η H_0.

Απορρίπτουμε την H_0, σε επίπεδο σημαντικότητας α, όταν $|t| > t_{n-2,\alpha/2}$.

Γνωρίζουμε όμως ότι για το μοντέλο της απλής γραμμικής παλινδρόμησης, ο ιδιος έλεγχος είναι $\mathrm{H}_0 : \beta_1 = 0$ προς $\mathrm{H}_1 : \beta_1 \neq 0$ και απορρίπτουμε αν $\left|t^*\right| > t_{n-2,\alpha/2}$, όπου $t^* = \dfrac{\hat{\beta}_1}{\sqrt{\widehat{Var}\left(\hat{\beta}_1\right)}}$.

Θα δείξουμε ότι οι δύο έλεγχοι είναι ισοδύναμοι, δηλαδή ότι $t = t^*$.

Παρατηρούμε ότι $R^2 = \dfrac{SS_{reg}}{SS_{tot}} = \dfrac{\hat{\beta}_1^{\,2}\cdot\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2}$, αλλά $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)}{\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2}$

$$\Rightarrow R^2 = \frac{\left[\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)\right]^2}{\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2\cdot\sum_{i=1}^{n}\left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2} = r^2.$$

Έχω επίσης ότι $\widehat{Var}\left(\hat{\beta}_1\right) = \dfrac{SS_{res}}{(n-2)\cdot\sum_{i=1}^{n}\left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2}$.

Άρα $t=\dfrac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}=\dfrac{\dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2}}}{\sqrt{\dfrac{1-R^2}{n-2}}}$.

Αλλά $1-R^2=1-\dfrac{SS_{reg}}{SS_{tot}}=\dfrac{SS_{tot}-SS_{reg}}{SS_{tot}}=\dfrac{SS_{res}}{SS_{tot}}$.

Οπότε:

$$t=\frac{\dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2}}}{\sqrt{\dfrac{SS_{res}}{(n-2)\cdot SS_{tot}}}}=\frac{\dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}}}{\sqrt{\dfrac{SS_{res}}{(n-2)}}}=$$

$$=\frac{\dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}}\cdot\dfrac{1}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}}}{\sqrt{\dfrac{SS_{res}}{(n-2)}}\cdot\dfrac{1}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}}}=\frac{\dfrac{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)}{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}}{\sqrt{\dfrac{SS_{res}}{(n-2)\cdot\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2}}}=\frac{\hat{\beta}_1}{\sqrt{\widehat{Var}\left(\hat{\beta}_1\right)}}=t^*$$

Θα αποδείξουμε ότι για το απλό γραμμικό μοντέλο, $\mathrm{Y}_i=\beta_0+\beta_1\cdot\mathrm{X}_i+\varepsilon_i$, ισχύει ότι $R^2=r^2\left(\mathrm{Y},\hat{\mathrm{Y}}\right)$, όπου

$$r\left(\mathrm{Y},\hat{\mathrm{Y}}\right)=\frac{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)\cdot\left(\hat{\mathrm{Y}}_i-\overline{\hat{\mathrm{Y}}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\hat{\mathrm{Y}}}\right)^2}}=\frac{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)\cdot\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2}}.$$

Έχουμε ότι $\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)\cdot\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)=\sum_{i=1}^{n}\left[\left(\mathrm{Y}_i-\hat{\mathrm{Y}}_i\right)+\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)\right]\cdot\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)=$

$$=\sum_{i=1}^{n}\left(\mathrm{Y}_i-\hat{\mathrm{Y}}_i\right)\cdot\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)+\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2=\sum_{i=1}^{n}e_i\cdot\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)+\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2=$$

$$=\sum_{i=1}^{n}e_i\cdot\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\cdot\sum_{i=1}^{n}e_i+\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2=\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2$$

Άρα $r^2\left(\mathrm{Y},\hat{\mathrm{Y}}\right)=\left[\dfrac{\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2}}\right]^2=\dfrac{\sum_{i=1}^{n}\left(\hat{\mathrm{Y}}_i-\overline{\mathrm{Y}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2}=R^2$.

Ο έλεγχος της υπόθεσης $\mathrm{H}_0:\rho=0$ προς $\mathrm{H}_1:\rho\neq 0$, μπορεί να γίνει και με το πηλίκο πιθανοφάνειας $\lambda\left(\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)=\dfrac{\sup_{\mathrm{H}_0} L\left(\underline{\theta}/\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)}{\sup_{\Theta} L\left(\underline{\theta}/\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)}$, όπου $\underline{\theta}=\left(\mu_{\mathrm{X}},\mu_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}},\rho\right)^{\mathrm{T}}$.

Απορρίπτουμε την H_0, αν $\lambda\left(\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)<c\Leftrightarrow\log\left[\lambda\left(\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)\right]<c^*$.

$$\left(\mathrm{X}_i,\mathrm{Y}_i\right)\sim\mathrm{N}_2\left(\begin{pmatrix}\mu_{\mathrm{X}}\\ \mu_{\mathrm{Y}}\end{pmatrix},\begin{pmatrix}\sigma_{\mathrm{X}}^2 & \rho\cdot\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}\\ \rho\cdot\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}} & \sigma_{\mathrm{Y}}^2\end{pmatrix}\right)$$

Άρα, $f\left(\underline{\theta}/\left(\mathrm{X}_i,\mathrm{Y}_i\right)\right)=\dfrac{1}{2\cdot\pi\cdot\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}\cdot\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2\cdot\left(1-\rho^2\right)}\cdot\left[\frac{(\mathrm{X}_i-\mu_{\mathrm{X}})^2}{\sigma_{\mathrm{X}}^2}+\frac{(\mathrm{Y}_i-\mu_{\mathrm{Y}})^2}{\sigma_{\mathrm{Y}}^2}-2\cdot\rho\cdot\frac{(\mathrm{X}_i-\mu_{\mathrm{X}})\cdot(\mathrm{Y}_i-\mu_{\mathrm{Y}})}{\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}}\right]}$

Οπότε, $L\left(\underline{\theta}/\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)=\prod_{i=1}^{n} f\left(\underline{\theta}/\left(\mathrm{X}_i,\mathrm{Y}_i\right)\right)=$

$$=\frac{1}{\left(2\cdot\pi\cdot\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}\cdot\sqrt{1-\rho^2}\right)^n}\cdot e^{-\frac{1}{2\cdot\left(1-\rho^2\right)}\cdot\left[\frac{\sum_{i=1}^{n}(\mathrm{X}_i-\mu_{\mathrm{X}})^2}{\sigma_{\mathrm{X}}^2}+\frac{\sum_{i=1}^{n}(\mathrm{Y}_i-\mu_{\mathrm{Y}})^2}{\sigma_{\mathrm{Y}}^2}-2\cdot\rho\cdot\frac{\sum_{i=1}^{n}(\mathrm{X}_i-\mu_{\mathrm{X}})\cdot(\mathrm{Y}_i-\mu_{\mathrm{Y}})}{\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}}\right]}$$

$$l\left(\underline{\theta}/\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)=\log\left[L\left(\underline{\theta}/\underline{\mathrm{X}},\underline{\mathrm{Y}}\right)\right]=-n\cdot\log\left(2\cdot\pi\cdot\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}\cdot\sqrt{1-\rho^2}\right)-$$
$$\frac{1}{2\cdot\left(1-\rho^2\right)}\cdot\left[\frac{\sum_{i=1}^{n}(\mathrm{X}_i-\mu_{\mathrm{X}})^2}{\sigma_{\mathrm{X}}^2}+\frac{\sum_{i=1}^{n}(\mathrm{Y}_i-\mu_{\mathrm{Y}})^2}{\sigma_{\mathrm{Y}}^2}-2\cdot\rho\cdot\frac{\sum_{i=1}^{n}(\mathrm{X}_i-\mu_{\mathrm{X}})\cdot(\mathrm{Y}_i-\mu_{\mathrm{Y}})}{\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}}\right]$$

Άρα, γενικά, παραγωγίζοντας την πιο πάνω παράσταση και θέτοντας την παράγωγο ίση με 0, έχουμε τις πιο κάτω εκτιμήτριες μεγίστης πιθανοθάνειας:

$$\hat{\mu}_{\mathrm{X}}=\overline{\mathrm{X}},\hat{\mu}_{\mathrm{Y}}=\overline{\mathrm{Y}},\hat{\sigma}_{\mathrm{X}}^2=\frac{1}{n}\cdot\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)^2,\hat{\sigma}_{\mathrm{Y}}^2=\frac{1}{n}\cdot\sum_{i=1}^{n}\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)^2,\hat{\rho}=r=\frac{\sum_{i=1}^{n}\left(\mathrm{X}_i-\overline{\mathrm{X}}\right)\cdot\left(\mathrm{Y}_i-\overline{\mathrm{Y}}\right)}{n\cdot\sigma_{\mathrm{X}}\cdot\sigma_{\mathrm{Y}}}$$

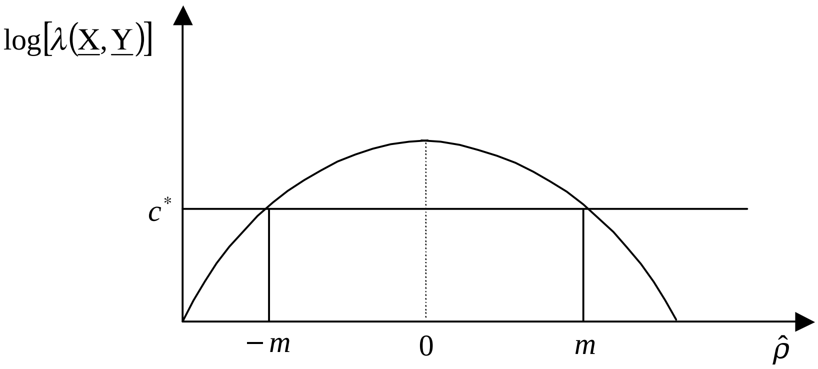
Όταν όμως ισχύει η $\mathrm{H}_0 : \rho = 0$, έχουμε τις εκτιμήτριες:

$$\hat{\hat{\mu}}_{\mathrm{X}} = \hat{\mu}_{\mathrm{X}} = \overline{\mathrm{X}}, \hat{\hat{\mu}}_{\mathrm{Y}} = \hat{\mu}_{\mathrm{Y}} = \overline{\mathrm{Y}}, \hat{\hat{\sigma}}_{\mathrm{X}}{}^2 = \hat{\sigma}_{\mathrm{X}}{}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathrm{X}_i - \overline{\mathrm{X}}\right)^2, \hat{\hat{\sigma}}_{\mathrm{Y}}{}^2 = \hat{\sigma}_{\mathrm{Y}}{}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathrm{Y}_i - \overline{\mathrm{Y}}\right)^2$$

Έχουμε ότι $\log\left[\lambda\left(\underline{\mathrm{X}}, \underline{\mathrm{Y}}\right)\right] = \log\left[L\left(\underline{\hat{\hat{\theta}}} \big/ \underline{\mathrm{X}}, \underline{\mathrm{Y}}\right)\right] - \log\left[L\left(\underline{\hat{\theta}} \big/ \underline{\mathrm{X}}, \underline{\mathrm{Y}}\right)\right] = \frac{n}{2} \cdot \log\left(1 - \hat{\rho}^2\right)$.

Άρα $\dfrac{d \log\left[\lambda\left(\underline{\mathrm{X}}, \underline{\mathrm{Y}}\right)\right]}{d\hat{\rho}} = -n \cdot \dfrac{\hat{\rho}}{1 - \hat{\rho}^2}$, $|\hat{\rho}| < 1$.

Έτσι έχουμε το παρακάτω γράφημα:



Άρα απορρίπτουμε την H_0, όταν $|\hat{\rho}| > m$.

<u>Μέση Τιμή και Πίνακας Συνδιακύμανσης Τυχαίου Διανύσματος</u>

<u>Ορισμός:</u>
Έστω $X_1, X_2, ..., X_n$ τυχαίες μεταβλητές.

Τότε το $(n \times 1)$ διάνυσμα $\underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_n \end{pmatrix}$ λέγεται n - διάστατο τυχαίο διάνυσμα.

Πιο γενικά μπορούμε να ορίσουμε τυχαίο πίνακα $\mathrm{X} = (X_{ij})$, όπου το (i, j) στοιχείο του είναι τυχαία μεταβλητή.

<u>Ορισμός:</u>

Η μέση τιμή του $\underline{X}$ ορίζεται από το διάνυσμα $\underline{\mu} = E(\underline{X}) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ . \\ . \\ . \\ E(X_n) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ . \\ \mu_n \end{pmatrix}$.

<u>Ορισμός:</u>
Ο πίνακας συνδιακύμανσης ενός $(n \times 1)$ τυχαίου διανύσματος $\underline{X}$ ορίζεται ως $\Sigma = Cov(\underline{X}) = E\left[(\underline{X} - \underline{\mu}) \cdot (\underline{X} - \underline{\mu})^T\right]$.

<u>Παρατηρήσεις:</u>

(1) Αν $n = 2$, τότε:

$$(\underline{X} - \underline{\mu}) \cdot (\underline{X} - \underline{\mu})^T = \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{pmatrix} \cdot (X_1 - \mu_1, X_2 - \mu_2) =$$

$$= \begin{bmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1) \cdot (X_2 - \mu_2) \\ (X_1 - \mu_1) \cdot (X_2 - \mu_2) & (X_2 - \mu_2)^2 \end{bmatrix}$$

Άρα $Cov(\underline{X}) = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var(X_2) \end{bmatrix}$.

Γενικότερα, αν το $\underline{X}$ είναι n- διάστατο, τότε:

$$Cov(\underline{X})=\begin{bmatrix} Var(X_1) & Cov(X_1,X_2) & Cov(X_1,X_3) & . & . & . & Cov(X_1,X_n) \\ Cov(X_2,X_1) & Var(X_2) & Cov(X_2,X_3) & . & . & . & Cov(X_2,X_n) \\ . & . & . & & & & . \\ . & . & . & & & & . \\ . & . & . & & & & . \\ Cov(X_n,X_1) & Cov(X_n,X_2) & Cov(X_n,X_3) & . & . & . & Var(X_n) \end{bmatrix}$$

Άρα ο πίνακας συνδιακύμανσης είναι τετραγωνικός, συμμετρικός, με διαγώνια στοιχεία τις διακυμάνσεις των X_i.

(2) Αν οι τυχαίες μεταβλητές X_i είναι ασυσχέτιστες, τότε ο $\Sigma = Cov(\underline{X})$ είναι διαγώνιος.

<u>Θεώρημα:</u>
Αν $\underline{X},\underline{Y}$ είναι $(n\times 1)$ τυχαία διανύσματα, $A_{(m\times n)}$ πίνακας, $\underline{b}_{(m\times 1)}$ διάνυσμα, $\underline{\alpha}(n\times 1)$ διάνυσμα και c πραγματική σταθερά, τότε ισχύουν:

(α) $E(\underline{X}\pm\underline{Y})=E(\underline{X})\pm E(\underline{Y})$

(β) $E(A\cdot\underline{X}\pm\underline{b})=A\cdot E(\underline{X})\pm\underline{b}$

(γ) $E(\underline{\alpha}^T\cdot\underline{X})=\underline{\alpha}^T\cdot E(\underline{X})$

(δ) $E(c\cdot\underline{X})=c\cdot E(\underline{X})$

(ε) $Cov(c\cdot\underline{X}\pm\underline{\alpha})=c^2\cdot Cov(\underline{X})$

(στ) Αν $\underline{X},\underline{Y}$ ασυσχέτιστα, τότε: $Cov(\underline{X}\pm\underline{Y})=Cov(\underline{X})\pm Cov(\underline{Y})$

(ζ) $Cov(A\cdot\underline{X})=A\cdot Cov(\underline{X})\cdot A^T$

(η) $Var(\underline{\alpha}^T\cdot\underline{X})=\underline{\alpha}^T\cdot Cov(\underline{X})\cdot\underline{\alpha}$

<u>Απόδειξη:</u>
Παραλείπεται.

<u>Θεώρημα:</u>
Έστω $\underline{X}_{(n\times 1)}$ τυχαίο διάνυσμα με $E(\underline{X})=\underline{\mu}$ και $Cov(\underline{X})=\Sigma$. Τότε:

(α) $\Sigma\geq 0$ (θετικά ημιορισμένος)

(β) $E\left(\|\underline{X}\|^2\right)=\|\underline{\mu}\|^2+tr(\Sigma)$, όπου $tr(\Sigma)$ είναι το ίχνος του πίνακα (το άθροισμα των διαγώνιων στοιχείων του) και $\|\underline{X}\|^2=\left(\underline{X}^T\cdot\underline{X}\right)=\sum_{i=1}^{n}X_i^{\ 2}$.

<u>Απόδειξη:</u>

(α) Έχουμε ότι για κάθε $\underline{\alpha} \in R^n$, ισχύει ότι $0 \le Var\left(\underline{\alpha}^{\mathrm{T}} \cdot \underline{\mathrm{X}}\right) = \underline{\alpha}^{\mathrm{T}} \cdot \Sigma \cdot \underline{\alpha} \Rightarrow$
$\Rightarrow \underline{\alpha}^{\mathrm{T}} \cdot \Sigma \cdot \underline{\alpha} \ge 0 \Rightarrow \Sigma \ge 0$.

(β) $\mathrm{E}\left(\left\|\underline{\mathrm{X}}\right\|^2\right) = \mathrm{E}\left(\sum_{i=1}^{n} \mathrm{X}_i^{\ 2}\right) = \sum_{i=1}^{n} \mathrm{E}\left(\mathrm{X}_i^{\ 2}\right) = \sum_{i=1}^{n} \left[\mathrm{E}^2(\mathrm{X}_i) + Var(\mathrm{X}_i)\right] =$
$= \sum_{i=1}^{n} \mathrm{E}^2(\mathrm{X}_i) + \sum_{i=1}^{n} Var(\mathrm{X}_i) = \sum_{i=1}^{n} \mu_i^{\ 2} + \sum_{i=1}^{n} Var(\mathrm{X}_i) = \left\|\underline{\mu}\right\|^2 + tr(\Sigma)$

Πλυδιάστατη Κανονική Κατανομή

Ορισμός:
Λέμε ότι το τυχαίο διάνυσμα $\underline{\mathrm{X}}$ ακολουθεί την πολυδιάστατη κανονική κατανομή, με μέση τιμή $\underline{\mu}$ και πίνακα διακυμάνσεων $\Sigma \ge 0$, αν

$$f_{\underline{\mathrm{X}}}(x) = \frac{1}{|\Sigma|^{1/2} \cdot (2 \cdot \pi)^{n/2}} \cdot e^{-\frac{1}{2}\cdot(\underline{x}-\underline{\mu})^{\mathrm{T}} \cdot \Sigma^{-1} \cdot (\underline{x}-\underline{\mu})}$$

. Συμβολικά, $\underline{\mathrm{X}} \sim \mathrm{N}_n(\underline{\mu}, \Sigma)$.

Ιδιότητες γραμμικών μετασχηματισμών:
(α) Αν $\underline{\mathrm{X}} \sim \mathrm{N}_n(\underline{\mu}, \Sigma)$, $\mathrm{A}_{(p\times n)}$ πίνακας και $\underline{b} \in R^n$, τότε:
$\mathrm{A} \cdot \underline{\mathrm{X}} \pm \underline{b} \sim \mathrm{N}_n\left(\mathrm{A} \cdot \underline{\mu} \pm \underline{b}, \mathrm{A} \cdot \Sigma \cdot \mathrm{A}^{\mathrm{T}}\right)$.
(β) Αν $\underline{\mathrm{X}} \sim \mathrm{N}_n(\underline{\mu}, \Sigma)$ και $c \in R$, τότε: $c \cdot \underline{\mathrm{X}} \sim \mathrm{N}_n\left(c \cdot \underline{\mu}, c^2 \cdot \Sigma\right)$.

Υποδιανύσματα, Ανεξαρτησία, Δεσμευμένες Κατανομές:
Έστω $\underline{\mathrm{X}} \sim \mathrm{N}_n(\underline{\mu}, \Sigma)$, $\underline{\mathrm{X}} = \begin{pmatrix} \underline{\mathrm{X}}_1 \\ \underline{\mathrm{X}}_2 \end{pmatrix}$, $\underline{\mu} = \begin{pmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{pmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, όπου $\underline{\mathrm{X}}_1, \underline{\mu}_1$ είναι p - διάστατα διανύσματα και Σ_{11} είναι $(p \times p)$ πίνακας. Τότε ισχύουν τα παρακάτω:
(1) $\underline{\mathrm{X}}_1 \sim \mathrm{N}_p\left(\underline{\mu}_1, \Sigma_{11}\right)$, $\underline{\mathrm{X}}_2 \sim \mathrm{N}_{n-p}\left(\underline{\mu}_2, \Sigma_{22}\right)$
(2) $\underline{\mathrm{X}}_1$ και $\underline{\mathrm{X}}_2$ είναι ανεξάρτητα $\Leftrightarrow$ $\Sigma_{12} = 0$
(3) Αν $\Sigma_{22} > 0$, τότε: $\underline{\mathrm{X}}_1 / \underline{\mathrm{X}}_2 = \underline{x}_2 \sim \mathrm{N}_p\left(\underline{v}_1, V_{11}\right)$, όπου
$\underline{v}_1 = \underline{\mu}_1 + \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \left(\underline{x}_2 - \underline{\mu}_2\right)$ και $V_{11} = \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21}$.

Απόδειξη για το (3):
Αναζητούμε την κατανομή του $\begin{pmatrix} \underline{\mathrm{Y}} \\ \underline{\mathrm{X}}_2 \end{pmatrix}$, όπου $\underline{\mathrm{Y}} = \underline{\mathrm{X}}_1 - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \underline{\mathrm{X}}_2$.

Αλλά, $\begin{pmatrix} \underline{Y} \\ \underline{X}_2 \end{pmatrix} = \begin{bmatrix} I & -\Sigma_{12} \cdot \Sigma_{22} \\ 0 & I \end{bmatrix} \cdot \begin{pmatrix} \underline{X}_1 \\ \underline{X}_2 \end{pmatrix}$.

Επειδή είναι γραμμικός μετασχηματισμός του $\underline{X}$, από την ιδιότητα (α) των γραμμικών μετασχηματισμών, έχουμε ότι:

$$\begin{pmatrix} \underline{Y} \\ \underline{X}_2 \end{pmatrix} \sim N_n\left(\begin{pmatrix} \underline{\mu}_1 - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \underline{\mu}_2 \\ \underline{\mu}_2 \end{pmatrix}, \begin{bmatrix} \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}\right)$$

Άρα τα $\underline{Y}$ και $\underline{X}_2$ είναι ανεξάρτητα και

$\underline{Y} \sim N_p\left(\underline{\mu}_1 - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \underline{\mu}_2, \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21}\right)$.

Αλλά, $\underline{Y} = \underline{X}_1 - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \underline{X}_2$.

Συνεπώς,

$$\left(\underline{X}_1 / \underline{X}_2 = \underline{x}_2\right) = \underline{Y} + \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \underline{x}_2 \sim N_p\left(\underline{\mu}_1 - \Sigma_{11} \cdot \Sigma_{22}^{-1} \cdot \left(\underline{x}_2 - \underline{\mu}_2\right), \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21}\right)$$

Παράδειγμα:

Έστω $n = 2$, $\underline{X} = \begin{pmatrix} \underline{X}_1 \\ \underline{X}_2 \end{pmatrix}$, $\underline{\mu} = \begin{pmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{pmatrix}$, $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho \cdot \sigma_1 \cdot \sigma_2 \\ \rho \cdot \sigma_1 \cdot \sigma_2 & \sigma_2^2 \end{bmatrix}$.

Επειδή $|\Sigma| = \sigma_1^2 \cdot \sigma_2^2 \cdot (1 - \rho^2)$, για να είναι θετικά ορισμένος ο Σ, έχουμε τις ακόλουθες συνθήκες: $\sigma_1^2 > 0, \sigma_2^2 > 0, |\rho| < 1$.

Οπότε, $\Sigma^{-1} = \dfrac{1}{1 - \rho^2} \cdot \begin{bmatrix} \dfrac{1}{\sigma_1^2} & -\dfrac{\rho}{\sigma_1 \cdot \sigma_2} \\ -\dfrac{\rho}{\sigma_1 \cdot \sigma_2} & \dfrac{1}{\sigma_2^2} \end{bmatrix}$.

Άρα η διδιάστατη κανονική, έχει συνάρτηση πυκνότητας πιθανότητας:

$$f(x_1, x_2) = \frac{1}{2 \cdot \pi \cdot \sigma_1 \cdot \sigma_2 \cdot \sqrt{1 - \rho^2}} \cdot e^{-\frac{1}{2 \cdot (1 - \rho^2)} \cdot \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} - 2 \cdot \rho \cdot \frac{(x_1 - \mu_1) \cdot (x_2 - \mu_2)}{\sigma_1 \cdot \sigma_2}\right]}$$

Επιπλέον, $X_1 / X_2 = x_2 \sim N\left(\mu_1 + \rho \cdot \dfrac{\sigma_1}{\sigma_2} \cdot (x_2 - \mu_2), \sigma_1^2 \cdot (1 - \rho^2)\right)$.

Τετραγωνικές Μορφές:

Αν $Z_1, Z_2, ..., Z_n \sim N(0,1)$, ανεξάρτητες $\Rightarrow \sum_{i=1}^{n} Z_i \sim X_n^2$.

Αλλά, $\underline{Z} = \begin{pmatrix} Z_1 \\ . \\ . \\ . \\ Z_n \end{pmatrix} \sim N_n(\underline{0}, I)$ και $\sum_{i=1}^{n} Z_i = \underline{Z}^T \cdot \underline{Z} = \|\underline{Z}\|^2$. Συνεπώς, $\underline{Z}^T \cdot \underline{Z} \sim X_n^2$.

<u>Θεώρημα:</u>
Αν $\underline{X} \sim N(\underline{\mu}, \Sigma), \Sigma > 0$, τότε $(\underline{X} - \underline{\mu})^T \cdot \Sigma^{-1} \cdot (\underline{X} - \underline{\mu}) \sim X_n^2$.

<u>Απόδειξη:</u>
Αφού $\Sigma > 0$, υπάρχει Σ^{-1} και $\Sigma^{-1/2} : \Sigma^{-1} = \Sigma^{-1/2} \cdot \Sigma^{-1/2}$.
Θεωρώ $\underline{Z} = \Sigma^{-1/2} \cdot (\underline{X} - \underline{\mu})$.
Τότε $\underline{Z} \sim N_n(\underline{0}, I) \Rightarrow \underline{Z}^T \cdot \underline{Z} \sim X_n^2$.
Συνεπώς, $\underline{Z}^T \cdot \underline{Z} = (\underline{X} - \underline{\mu})^T \cdot \Sigma^{-1/2} \cdot \Sigma^{-1/2} \cdot (\underline{X} - \underline{\mu}) = (\underline{X} - \underline{\mu})^T \cdot \Sigma^{-1} \cdot (\underline{X} - \underline{\mu})$