3) **Kaj veš o prvi in drugi regresijski premici (npr. kje se sečeta) ter o metodi najmanjših kvadratov (definiraj tudi kovarianco in (Pearsonov) koeficient korelacije za dve številski spremenljivki ter pojasni njegov pomen za regresijo)?**

Kovarianca je $Cov(X,Y) = E((E(X)-X)(E(Y)-Y)) = E(XY) - E(X)E(Y)$, in jo uporabljamo za ugotavljanje zveze med dvema slučajnima spremenljivkama.

Za **Pearsonov koeficient korelacije** $r(X,Y) = \dfrac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$, vemo, da leži na intervalu $[-1,1]$. Z njim lahko ugotovimo, če sta spremenljivki močno linearno povezani, saj nam enakost, ko velja $|Cov(X,Y)| = \sigma(X)\sigma(Y)$ zagotavlja prav to.

Vrednost $r(X,Y)=1$ predstavlja popolno pozitivno povezanost spremenljivk, $r(X,Y)=-1$ predstavlja popolno negativno povezanost spremenljivk, pri $r(X,Y)=0$ (tj. $E(XY)=E(X)E(Y)$ oziroma $D(X+Y)=D(X)+D(Y)$) pa pravimo, da sta spremenljivki nekorelirani (velja, da sta neodvisni spremenljivki vedno nekorelirani, nekorelirani spremenljivki pa nista nujno neodvisni).

Ko ugotovimo, da med spremenljivkama obstaja močna linearna odvisnost, lahko le-to poiščemo z metodo najmanjših kvadratov. Kar pomeni, da funkcijo $f(a,b) = \sum_{i=1}^{n} (y_i - (bx_i + a))^2$ parcialno odvajamo po a in b ter dobimo dve enačbi iz katerih določimo a in b. Na ta način pridemo do regresijskih premic $Y = E(Y) + \dfrac{Cov(X,Y)}{\sigma_x^2}(X - E(X))$ ter $X = E(X) + \dfrac{Cov(X,Y)}{\sigma_y^2}(Y - E(Y))$, ki se sečeta v točki $(E(X), E(Y))$.