Dagens tekst

Kap 4: Matematisk forventning

Viktige omgrep

- Forventningsverdi
- Varians / standard avvik
- Kovarians / korrelasjon

Viktige reknereglar

Forventning til lineærkombinasjon:

$$E(a_0+a_1X_1+\ldots a_nX_n) = a_0+a_1E(X_1)+\cdots+a_nE(X_n)$$

Varians til lineærkombinasjon:

Forventningsverdi

Definisjon 4.1

La X vere ein stok.var. med sannsynsfordeling $f(x)$. Forventningsverdien til X er då

$$\mu = E(X) = \sum_{\forall x} x f(x)$$

dersom X er diskret, og

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

dersom X er kontinuerleg.

Tolkning:

- Gjennomsnitt av uendeleg mange data trukke frå $f(x)$
- Tyngdepunktet i fordelinga.

Forventningsverdi av funksjon av stokastisk variabel

Teorem 4.1

For ein funksjon g(X);

$$\mu_{g(X)} = E(g(X)) = \sum_x g(x)f(x)$$

for diskret X, og

$$\mu_{g(X)} = E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

for kontinuerleg X.

Varians

Definisjon 4.3

La X vere ein stok. var. med forventning μ. Variansen til X er då:
For diskret X:

$$Var(X) = E((X-\mu)^2) = \sum_{\forall x} (x-\mu)^2 f(x)$$

For kontinuerleg X:

$$Var(X) = E((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$

Er interessert i kostnaden på leigebil i ferien.
Pris:

- Fast pris: $a = 500$ €
- Km pris: $b = 0.2$ €/ km

Her ei sannsynsfordeling $f(x)$ på lengda X som skal kjørast.

- $E(X) = 2000 km$
- $Var(X) = (500 km)^2$

Kostnad

$$Y = a + bX$$

Forventning og varians til lineær funksjon

Teorem 4.5 og 4.9, VIKTIG

La X vere ein stok. var. med $E(X) = \mu_X$ og $Var(X) = \sigma_X^2$, og la $Y = a + bX$, der a og b er kjende konstantar. Då er;

$$E(Y) = a + b\mu_X$$

$$Var(Y) = b^2\sigma_X^2$$

Definisjon 4.4

La X og Y vere stokastiske variable med simultan sannsynsfordeling $f(x,y)$, og forventning hhv μ_x og μ_y.
Kovariansen til X og Y er då:

$$Cov(X,Y) = \sigma_{xy} = E[(X-\mu_X)(Y-\mu_Y)]$$

For kontinuerleg X og Y:

$$Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x,y)dxdy$$

For diskret X og Y:

$$Cov(X,Y) = \sum_{\forall x}\sum_{\forall y}(x-\mu_X)(y-\mu_Y)f(x,y)$$

Eksempel, berehjelp for mormor

Eg er med mormor på butikken som berehjep. Hor mykje må eg bere?

Ho kjøper:

- alltid 2 liter melk ($a = 2kg$)
- X pakkar mjøl, kvar på $b = 2kg$
- Y pakkar sukker, kvar på $c = 1kg$

Veit at:

- $E(X) = \mu_X = 1.5$ og $Var(X) = \sigma_X^2$
- $E(Y) = \mu_Y = 2$ og $Var(Y) = \sigma_Y^2$

Totalt å bere:

$$Z = a + bX + cY$$

$$E(Z) = a + b\mu_X + c\mu_Y$$

Forventning og varians av lineærkombinasjon

Teorem, VIKTIG

Dersom $a_0, a_1, \ldots, a_n$ er konstantar og $X_1, X_2, \ldots X_n$ stok. var., så er

$$E(a_0 + \sum_{i=1}^{n} a_i X_i) = a_0 + \sum_{i=1}^{n} a_i E(X_i)$$

og

$$Var(a_0 + \sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} a_i a_j Cov(X_i, X_j)$$

for *uavhengige* $X_1, \ldots, X_n$ er

$$Var(a_0 + \sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i)$$

Kun for lineær kombinasjonar....

fristelsar.....

$Z = X/Y$

$$E(Z) \neq \frac{E(X)}{E(Y)}$$

...only the hard way...

Må bruke definisjonen på forventning for funksjon av to stok.var. Kan evt. approksimere vha Taylorrekker.

Chebychevs teorem

Teorem 4.11

Sannsynet for at ein stokastisk variabel tar ein verdi innanfor k standardavvik frå forventningsverdien er minst $1 - 1/k^2$.
Dvs

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

Viktigaste resultat

- $E(a + bX) = a + bE(X)$
- $Var(a + bX) = b^2 Var(X)$
- $E(a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_n X_n) = a_0 + a_1 E(X_1) + a_2 E(X_2) + \ldots a_n E(X_n)$

Dersom $X_1, X_2, \ldots X_n$ er uavhengige

- $Var(a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_n X_n) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + \ldots a_n^2 Var(X_n)$

Viktige omgrep

- Forventningsverdi
- Varians / standard avvik
- Kovarians / korrelasjon

Viktige rekneregler

Forventning til lineærkombinasjon:

$$E(a_0+a_1X_1+\dots a_nX_n) = a_0+a_1E(X_1)+\dots+a_nE(X_n)$$

Varians til lineærkombinasjon: **Dersom $X_1, X_2, \dots X_n$ er uavhengige**

$$Var(a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n) = a_1^2Var(X_1) + a_2^2Var(X_2) + \dots a_n^2Var(X_n)$$