Repetition (Urplock)

V1: Utfallsrum, slumphändelse. Divisionsregeln:

$$P(A) = \frac{\#\text{ gynnsamma utfall}}{\#\text{ utfall}}$$

~~Bayes lag:~~ Bet. sann.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

A, B oberoende om $P(A \cap B) = P(A)\,P(B)$

V2: Diskreta s.v. pmf: $\{p_k\}$ $p_k = P(X = x_k)$

cdf: $F(x) = P(X \leq x)$ defin. enkla egenskaper.

Vi def. Väntevärde $E[X]$ och varians $Var(X)$

Fördelningar. 1) Likformig 2) $Bern(p)$ 3) $Bin(n,p)$
4) $Hg(N,n,p)$ 5) $Geom(p)$ 6) $Nb(r,p)$
7) $Poi(\lambda)$

V3: Kont. s.v. Vi def cdf $F(x) = P(X \leq x)$

och pdf: $f(x) = F'(x)$. Vi def. $\underbrace{E[g(X)]}_{= \int g(x) f(x)\,dx}$, $Var(X)$.

Fördelningar: 1) ~~B~~ $U(a,b)$ 2) $Exp(\lambda)$ 3) $Gamma(\alpha, \lambda)$
4) $N(\mu, \sigma^2)$

Transformation av s.v. och egenskaper för dessa.

V4: Gemensamma fördelningar, marginal och beting. -förd.

Diskret: 1) $p_{XY}(n,k) = P(X = n, Y = k)$ G-f.

2) $p_X(n) = P(X = n) = \sum_k P(X = n, Y = k)$ m-f (för X)

3) $\mathbb{P}_{Y|X}(k|n) = \dfrac{\mathbb{P}_{XY}(n,k)}{\mathbb{P}_X(n)} = \dfrac{P(X=n, Y=k)}{P(X=n)}$ bet.-förd.

X, Y oberoende om $p_{XY}(n,k) = p_X(n)\, p_Y(k) \quad \forall\ (n,k)$

Kont.: 1)

$$P(a<X<b,\ c<Y<d) = \int_a^b \int_c^d \underbrace{f_{XY}(x,y)\,dx\,dy}_{\text{gen. pdf.}}$$

2)

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy \quad \text{– marg.-förd. } (X)$$

3)

Bet. pdf: $f_{Y|X}(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$

X, Y oberoende om $f_{XY}(x,y) = f_X(x)\, f_Y(y)$

Används t.ex. för $E[X|Y]$ och $Var(X|Y)$

Vi def. Covarians och Correlation.

Fördelniger: 1) $Mn(n: p_1, \dots, p_r)$ 2) $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ (biv. norm.)

Momentgenererande funktion:

Def: $M(t) = E[e^{tX}]$ om den $\exists$.

1) $X \stackrel{d}{=} Y \iff M_X(t) = M_Y(t) \quad \forall\ t \in (-\varepsilon, \varepsilon)$

2) $E[X^n] = M^{(n)}(0), \quad \forall\ n \in \mathbb{N}$

3) $E[e^{tX + sY}] = E[e^{tX}]\, E[e^{sY}] \iff X, Y$ ober.

4) $X_n \to X \iff M_{X_n}(t) \to M_X(t) \quad \forall\ t \in (-\varepsilon, \varepsilon)$

V5:

LLN: $\bar{X}_n := \frac{X_1 + \dots + X_n}{n}$ $E[X_i] = \mu$, $Var(X_i) = \sigma^2$

och X_i ober. $\Rightarrow$ $\varepsilon > 0$.

$$P(|\bar{X}_n - \mu| > \varepsilon) \to 0, \quad n \to \infty$$

CLT: Låt $S_n = X_1 + \dots + X_n$ där X_i ober. med

$E[X_i] = \mu$, $Var(X_i) = \sigma^2$

$$\Rightarrow \quad P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) \xrightarrow[n\to\infty]{} \Phi(x) \quad \forall x$$

Approx: 1) $Bin(n, p) \approx N(np, npq)$ $np, nq \ge 5$

2) $Hg(N, n, p) \approx N\left(np, npq\frac{N-n}{N-1}\right)$, $np, nq \ge 5$

o.s.v.

$$\chi^2_k \overset{d}{:=} Z_1^2 + \dots + Z_k^2 \qquad Z_i \sim N(0,1) \text{ iid}$$

chi-kvadrat med k frihetsgrader

Vi har: $\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$ $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

X_i ober. $\sim N(\mu, \sigma^2)$

V6: t-förd $Z \sim N(0,1)$ $X \sim \chi^2_k$

$$\Rightarrow \frac{Z\sqrt{k}}{\sqrt{X}} \sim t_k\text{-förd.}$$

$$\Rightarrow \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1} \quad \text{(vanlig test-statistika)}$$

Skattningar: $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ skattar θ.

Hur väljer vi $\hat{\theta}$? 1) uppenbart 2) MME 3) MLE

4) korrigerad MLE

V6 forts: Väntevärdesriktig skattare:

$E[\hat{\theta}] = \theta$.

Konsistent skattare: $E[(\hat{\theta}-\theta)^2] \xrightarrow[n \to \infty]{} 0$

$\sigma_{\hat{\theta}}$ kallas för standardfelet. $s_{\hat{\theta}}$ är skattaren av $\sigma_{\hat{\theta}}$.

Konfidensintervall: Skattar θ mha $\hat{\theta}$ och skapar ett $100(1-\alpha)$ % intervall runt $\hat{\theta}$.

Vi använder ofta Normalapproximation (CLT).

Ibland måste vi räkna med den exakta fördelningen.

V7: Hypotestest:

Vi ställer två hypoteser mot varandra.

α = signifikansnivån = P(förk. $H_0 \mid H_0$)

$1-\beta$ = teststyrkan = P(förk. $H_0 \mid H_1$)

Test-statistika: En funktion av datamängden med olika värden under H_0 och H_1.

Förkastningsområde $RR = \{T :$ vi förk. $H_0\}$

Likelihood-ratio test: $\dfrac{f(\underline{X} \mid H_0)}{f(\underline{X} \mid H_1)}$

Gener. likelihood-ratio test: $\dfrac{\max_{w \in \Omega_0} f(\underline{X} \mid w)}{\max_{w \in \Omega} f(\underline{X} \mid w)}$