

# 9MA241/341: Formelblad

Johan Thim\*

27 maj 2012

## Kombinatorik

Om  $n$  är ett icke-negativt heltal definieras  $n!$  som  $n \cdot (n-1) \cdot (n-2) \cdots 1$  om  $n \geq 1$  och  $0! = 1$ . Talet  $n!$  beskriver på hur många sätt  $n$  distinkta objekt kan ordnas (antalet permutationer).

Låt  $k$  och  $n$  vara icke-negativa heltal så att  $k \leq n$ . Binomialkoefficienten  $n$  över  $k$  definieras som

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

Att välja  $k$  element av  $n$  kan göras på följande antal sätt:

	Med återläggning	Utan återläggning
Med ordning	$n^k$	$\frac{n!}{(n-k)!}$
Utan ordning	$\frac{(n+k-1)!}{k!(n-1)!}$	$\frac{n!}{(n-k)!k!}$

## Sannolikhetslära

Låt  $A$  och  $B$  vara händelser i ett utfallsrum  $\Omega$  med sannolikhetsmått  $P$ . Följande samband gäller:

1. Komplementhändelse:

$$P(A^*) = 1 - P(A).$$

2. Unionen:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

3. Händelserna  $A$  och  $B$  är oberoende om och endast om

$$P(A \cap B) = P(A)P(B).$$

4. Om  $P(A) > 0$  ges den betingade sannolikheten för  $B$  (betingat på att  $A$  inträffar) av

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

5. Lagen om total sannolikhet (LOTS): Låt  $H_1, H_2, \dots, H_n$  vara parvis disjunkta händelser ( $H_i \cap H_j = \emptyset$  om  $i \neq j$ ) så att  $H_1 \cup H_2 \cup \dots \cup H_n = \Omega$ . Då gäller att

$$P(A) = \sum_{i=1}^n P(A|H_i)P(H_i).$$

6. Bayes sats: med samma förutsättningar som för LOTS gäller att

$$P(H_i|A) = \frac{P(A|H_i)P(H_i)}{\sum_{k=1}^n P(A|H_k)P(H_k)}.$$

\*jothi@mai.liu.se

## Väntevärde och varians

För en slumpvariabel  $X$  och en funktion  $g(x)$  gäller att

$$E(g(X)) = \sum_k g(k)p_X(k)$$

om  $X$  är diskret och

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

om  $X$  är kontinuerlig. Variansen fås genom

$$V(X) = E((X - E(X))^2) = E(X^2) - E(X)^2.$$

Den sista likheten brukar kallas Steiners sats.

Om  $X_1, X_2, \dots, X_n$  är stokastiska variabler och  $a_1, a_2, \dots, a_n, b$  är reella konstanter så gäller att

$$E\left(b + \sum_{k=1}^n a_k X_k\right) = b + \sum_{k=1}^n a_k E(X_k).$$

Om variablerna dessutom är oberoende så gäller att

$$V\left(b + \sum_{k=1}^n a_k X_k\right) = \sum_{k=1}^n a_k^2 V(X_k).$$

Kovariansen  $C(X, Y)$  mellan två stokastiska variabler kan beräknas enligt

$$C(X, Y) = E(XY) - E(X)E(Y).$$

## Vanliga fördelningar

Låt  $X$  vara en stokastisk variabel och  $P$  sannolikhetsmåttet. Fördelningsfunktionen  $F(x)$  definieras som

$$F(x) = P(X \leq x) \text{ för alla } x \in \mathbb{R}.$$

### Diskreta fördelningar

Sannolikhetsfunktion:  $p_X(k) = P(X = k)$ .

- Binomialfördelning:  $X \sim \text{Bin}(n, p)$ ,  $n \geq 0$  heltal,  $0 < p < 1$ ,

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n,$$

$$E(X) = np, \quad V(X) = np(1-p).$$

- Hypergeometrisk:  $X \sim \text{Hyp}(N, n, p)$ ,  $0 < p < 1$ ,

$$p_X(k) = \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}}, \quad 0 \leq k \leq Np, \quad 0 < n-k \leq N(1-p),$$

$$E(X) = np, \quad V(X) = \frac{N-n}{N-1} np(1-p).$$

- Ffg:  $X \sim \text{Ffg}(p)$ ,  $0 < p < 1$ ,

$$p_X(k) = p(1-p)^{k-1}, \quad k = 1, 2, 3, \dots,$$

$$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2}.$$

- Poisson:  $X \sim \text{Po}(\mu)$ ,  $\mu > 0$ ,

$$p_X(k) = \frac{\mu^k}{k!} e^{-\mu}, \quad k = 0, 1, 2, \dots,$$

$$E(X) = \mu, \quad V(X) = \mu.$$

## Kontinuerliga fördelningar

Täthetsfunktionen  $f_X(x)$  uppfyller:

$$P(X \leq x) = \int_{-\infty}^x f(s) ds, \quad \text{för alla } x \in \mathbb{R}.$$

- Rektangelfördelning (likformig):  $X \sim R(a, b)$ , där  $a < b$ ,

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{för övrigt,} \end{cases}$$

$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}.$$

- Normalfördelning:  $X \sim N(\mu, \sigma)$ ,  $\mu \in \mathbb{R}$  och  $\sigma > 0$ ,

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

$$E(X) = \mu, \quad V(X) = \sigma^2.$$

- Exponentialfördelning:  $X \sim \text{Exp}(\mu)$ ,  $\mu > 0$ ,

$$f_X(x) = \begin{cases} \frac{1}{\mu} \exp(-x/\mu), & x \geq 0, \\ 0, & \text{för övrigt,} \end{cases}$$

$$E(X) = \mu, \quad V(X) = \mu^2.$$

Ibland används intensiteten  $\lambda = 1/\mu$  som parameter istället.

## Centrala gränsvärdessatsen

Låt  $X_1, X_2, X_3, \dots$  vara en följd av oberoende och likafördelade stokastiska variabler med  $E(X_k) = \mu$  och  $V(X_k) = \sigma^2$  för alla  $k = 1, 2, 3, \dots$ . Då gäller att

$$X = \sum_{k=1}^n X_k \stackrel{\text{appr.}}{\sim} N(n\mu, \sigma\sqrt{n})$$

för stora  $n$  i den meningen att

$$\lim_{n \rightarrow \infty} P\left(a < \frac{X - n\mu}{\sigma\sqrt{n}} < b\right) = \Phi(b) - \Phi(a)$$

för alla  $a, b \in \mathbb{R}$ . För god approximation brukar  $n \geq 30$  duga.

## Approximationer

$$\text{Hyp}(N, n, p) \stackrel{\text{appr.}}{\sim} \text{Bin}(n, p) \text{ om } \frac{n}{N} \leq \frac{1}{10}.$$

$$\text{Hyp}(N, n, p) \stackrel{\text{appr.}}{\sim} N\left(np, \sqrt{\frac{N-n}{N-1}np(1-p)}\right) \text{ om } \frac{N-n}{N-1}np(1-p) \geq 10.$$

$$\text{Bin}(n, p) \stackrel{\text{appr.}}{\sim} \text{Po}(np) \text{ om } n \geq 10 \text{ och } p \leq 0.1.$$

$$\text{Bin}(n, p) \stackrel{\text{appr.}}{\sim} N(np, \sqrt{np(1-p)}) \text{ om } np(1-p) \geq 10.$$

$$\text{Po}(\mu) \stackrel{\text{appr.}}{\sim} N(\mu, \sqrt{\mu}) \text{ om } \mu \geq 15.$$

## Statistik

### Stickprov

Vi kallar  $X_1, X_2, \dots, X_n$  för ett stickprov (av storlek  $n$ ) av en stokastisk variabel  $X$  om alla  $X_k$  är oberoende av varandra och har samma fördelning som  $X$ . Ibland kallar vi stickprovet för ett stickprov av fördelningen för  $X$ . Medelvärde och stickprovsvariansen,

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k \quad \text{och} \quad S^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2,$$

är väntevärdesriktiga och konsistenta skattningar av väntevärdet respektive variansen för fördelningen. För att skatta standardavvikelsen brukar  $S = \sqrt{S^2}$  användas.

Om  $X \sim N(\mu, \sigma)$  så gäller att

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \quad \text{och} \quad \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1).$$

### Normalapproximation

Om  $X \sim \text{Bin}(n, p) \stackrel{\text{appr.}}{\sim} N(np, \sqrt{np(1-p)})$ , så gäller med  $\hat{p} = X/n$  att

$$\frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} \stackrel{\text{appr.}}{\sim} N(0, 1).$$

Om  $X \sim \text{Po}(\mu) \stackrel{\text{appr.}}{\sim} N(\mu, \sqrt{\mu})$ , så gäller med  $\hat{\mu} = X$  att

$$\frac{\hat{\mu} - \mu}{\sqrt{\hat{\mu}}} \stackrel{\text{appr.}}{\sim} N(0, 1).$$

### Två stickprov

Låt  $X_1, X_2, \dots, X_n$  och  $Y_1, Y_2, \dots, Y_m$  vara två av varandra oberoende stickprov från  $N(\mu_X, \sigma)$  respektive  $N(\mu_Y, \sigma)$  (samma  $\sigma$ ). Om  $S_X^2$  och  $S_Y^2$  är de respektive stickprovsvarianserna så ges den sammanvägda skattningen av  $\sigma^2$  av

$$S^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{(n-1) + (m-1)}.$$

Om  $a$  och  $b$  är reella tal så är

$$\frac{(a\bar{X} + b\bar{Y}) - (a\mu_X + b\mu_Y)}{\sigma\sqrt{\frac{a^2}{n} + \frac{b^2}{m}}} \sim N(0, 1) \quad \text{och} \quad \frac{(a\bar{X} + b\bar{Y}) - (a\mu_X + b\mu_Y)}{S\sqrt{\frac{a^2}{n} + \frac{b^2}{m}}} \sim t(n+m-2).$$

### Enkel linjär regression

Om  $y_j$  är observationer av stokastiska variabler  $Y_j = b_0 + b_1x_j + \epsilon_j$  för  $j = 1, 2, \dots, n$ , där  $\epsilon_j \sim N(0, \sigma)$  är oberoende och  $(x_j, y_j)$  är uppmätta datapunkter, så kan de värden på  $b_0$  och  $b_1$  som minimerar

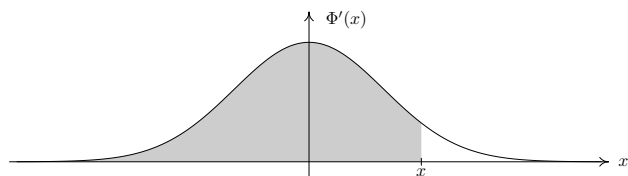
$$\sum_{j=1}^n (y_j - (b_0 + b_1x_j))^2$$

beräknas enligt följande:

$$b_0 = \bar{y} - b_1\bar{x} \quad \text{och} \quad b_1 = \frac{\sum_{j=1}^n x_j y_j - n\bar{x}\bar{y}}{\sum_{j=1}^n x_j^2 - n\bar{x}^2}.$$

## Normalfördelning: $X \sim N(0,1)$

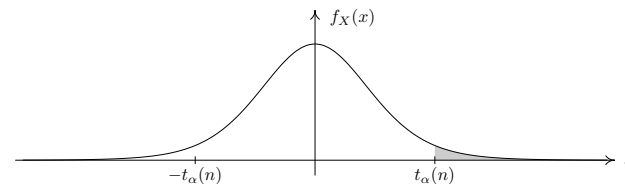
I tabellen nedan finns värden för  $\Phi(x) = P(X \leq x)$ , där  $X \sim N(0,1)$ , tabulerade. Om  $x < 0$ , utnyttja att  $\Phi(x) = 1 - \Phi(-x)$ .



x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

## t-fördelning: $X \sim t(n)$

Nedan finns  $\alpha$ -kvantiler för  $t$ -fördelningen, d v s värden  $t_\alpha(n)$  så att  $P(X > t_\alpha(n)) = \alpha$  om  $X \sim t(n)$ .



n	$\alpha$						
	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	1.291	1.662	1.987	2.368	2.632	3.183	3.402
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
$\infty$	1.282	1.645	1.960	2.326	2.576	3.090	3.291

