Lösningar till tentamen i Matematisk statistik och diskret matematik D2 (MVE055/MSG810).
Den 23 oktober 2007.

1. Lösning:
Låt X vara antal flickor bland de $n = 105913$ barnen. Då är $X \sim \mathrm{Bin}(n, p)$. Både X och $\hat{p} = \frac{X}{n}$ är approximativt normalfördelade vilket ger det 95%-iga konfidensintervallet:
$$\hat{p} \pm z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
Med våra värden fås $\hat{p} = \frac{51430}{105913} \approx 0.486$ och konfidensintervallet
$$0.486 \pm 1.96 \cdot 0.00154 \approx 48.6\% \pm 0.2\%$$

2. Lösning:
Låt X vara antal kort som behöver dras. X kan anta värdena 2 eller 3. Beteckna utfallen med en sekvens av b (blå) eller s (svart) i den ordning korten drogs.
$$\begin{aligned} \mathrm{P}(X=2) &= \mathrm{P}(\{bb, ss\}) = \mathrm{P}(\{bb\}) + \mathrm{P}(\{ss\}) = \frac{2}{6}\cdot\frac{1}{5} + \frac{4}{6}\cdot\frac{3}{5} = \frac{2+12}{30} = \frac{7}{15} \\ \mathrm{P}(X=3) &= 1 - \mathrm{P}(X=2) = \frac{8}{15} \\ \mathrm{E}[X] &= \frac{7}{15}\cdot 2 + \frac{8}{15}\cdot 3 = \frac{38}{15} = 2.5333\ldots \end{aligned}$$

3. Lösning:
   a) T.ex. $B = \{1, 3, 5\}$
   b) T.ex. $B = \{3, 4, 5, 6\}$
   c) T.ex. $B = \{1, 2, 3, 4, 5\}$

4. Lösning:
   a) Konfidensintervallet ges av $\hat{x} \pm t_{0.05}\frac{s}{\sqrt{10}} = -3.12 \pm 0.09$, där $t_{0.05} \approx 1.833$ fås från t-fördelningen med 9 frihetsgrader.
   b) Av intervallet ovan fås: $[2^{-3.21}, 2^{-3.03}] \approx [0.108, 0.123]$
   c) Konfidensintervallet för variansen ges av $\left[\frac{9s^2}{\chi_{0.05}}, \frac{9s^2}{\chi_{0.95}}\right] \approx [0.05305, 0.01043]$, varav konfidensintervallet för standardavvikelsen blir $[0.10, 0.23]$

5. Lösning:
$m_x'(t) = \frac{e^t + 2e^{2t}}{4}$; $m_x''(t) = \frac{e^t + 4e^{2t}}{4}$
$\mathrm{Var}\, X = E[X^2] - (E[X])^2 = m_x''(0) - (m_x'(0))^2 = \frac{5}{4} - \left(\frac{3}{4}\right)^2 = \frac{11}{16}$

6. Lösning:
Låt X vara antal vinster. Då är $X \sim \mathrm{Bin}(1000, \frac{1}{2})$. Enligt central gränsvärdessatsen gäller har X approximativt samma fördelning som $Z \sim \mathrm{Norm}(500, \sqrt{250})$. Den sökta sannolikheten är:

$$\mathrm{P}(|X-500|>50) \approx \mathrm{P}(|Z-500|>50.5) = \mathrm{P}\left(\frac{|Z-500|}{\sqrt{250}} > \frac{50.5}{\sqrt{250}}\right) =$$
$$= 2\left(1-\phi\left(\frac{50.5}{\sqrt{250}}\right)\right) \approx 2(1-\phi(3.19)) \approx 2(1-0.99929) \approx 0.14\%$$

7. Lösning:

$$\mathrm{Cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = \mathrm{Var}\,X$$

8. Lösning:

a) Låt X vara antal mail som inkommer mellan kl 14.00 och 15.00. Då är $X \sim \mathrm{Po}(5)$ och $\mathrm{P}(X=0) = \frac{e^{-5}\cdot 5^0}{0!} = e^{-5} \approx 0.7\%$

b) $Y \sim \mathrm{Geom}(0.01)$ ger $\mathrm{E}[Y] = \frac{1}{0.01} = 100$ och $\sigma_Y = \frac{\sqrt{1-0.01}}{0.01} \approx 99.5$

9. Lösning:
X: din hårddisks förväntade livslängd, Y: din hårddisks faktiska livslängd
Då är $X = \begin{cases} 30 & m.s.\ \ 0.8 \\ 2 & m.s.\ \ 0.2 \end{cases}$ och $Y \sim \mathrm{Exp}(X)$ Bayes sats ger nu

$$\mathrm{P}(X=2|Y\geq 3) = \frac{\mathrm{P}(Y\geq 3|X=2)\,\mathrm{P}(X=2)}{\mathrm{P}(Y\geq 3|X=2)\,\mathrm{P}(X=2) + \mathrm{P}(Y\geq 3|X=30)\,\mathrm{P}(X=30)} =$$
$$= \frac{e^{-3/2}\cdot 0.2}{e^{-3/2}\cdot 0.2 + e^{-3/30}\cdot 0.8} \approx 5.8\%$$