Lösningar till tentamen i Matematisk statistik och diskret matematik D2 (MVE055/MSG810).
Den 19 oktober 2010. These are sketches of the solutions.

1. Lösning:
   a) $P(A \cap B) = P(A)P(B)$, the knowledge about one event does not carry any information about the other
   b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.6 + 0.5 - 0.6 \cdot 0.5 = 0.8$
   $P(A \cap B^C) = P(A) - P(A \cap B) = 0.6 - 0.3 = 0.3$
   c) $P(G|H) := P(G \cap H)/P(H)$. It A and B independent then $P(A|B) = P(A)$.

2. Lösning:
   a) 0.1
   b) 0.15
   c) $P(X_1 + X_2 \in \{2, 4\}) = P(\{1 + 1\}) + P(\{1 + 3\}) + P(\{3 + 1\}) + P(\{2 + 2\}) = 0.3225$

3. Lösning:
   a) $F_X(x) = P(X \le x)$, $f_X = \frac{\mathrm{d}}{\mathrm{dx}} F_X(x)$
   b) $F_X(x) = \int_{x_0}^{x} Cy^{-\alpha}\mathrm{dy} = C\frac{1}{1-\alpha}(x^{1-\alpha} - x_0^{1-\alpha})$, we need to choose some lower bound as there is the problem of integrability at 0

4. Lösning:
   a) $f_X(x) = \lambda e^{-\lambda x}$ $x \ge 0$ $\mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\mathrm{dx} = \frac{1}{\lambda}$.
   b) $3 \cdot 1 - 7 \cdot 1 = -4$
   c) $9 \cdot 1 + 49 \cdot 1 = 58$, we use that if X and Y independent then
   $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ and $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$. $\mathrm{Cov}[X, Y] = 0$
   d) If the covariance is non–zero then they are dependent. A zero covariance does not allow us to draw conclusions except in the case of a normal distribution where zero covariance implies independence.

5. Lösning:
   a) $\hat{\theta}$ is unbiased if $\mathbf{E}[\hat{\theta}] = \theta$, $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ $\mathbf{E}[\overline{X}] = \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}[X_i] = \mu$, therefore is unbiased, $\overline{X} \sim \mathcal{N}(\mu, 0.5)$, $\overline{X}$ is normal as linear combination of normals then we need to calculate its mean and variance using the properties of the mean and variance of a linear combination of independent random variables
   b) Motivate $P(-z_{\alpha/2} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le z_{a/2}) = 1 - \alpha$ and transform it to $\mu \in \overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$. If we would have a large number of independent samples of size n then about $(1 - \alpha)100\%$ of them would generate confidence intervals which contain the true value of μ.
   c) 42.88312 ± 0.98 not surprised as 42.2 is inside confidence interval

6. Lösning:
   You need to calculate S_j^2 as $S_j^2 = \frac{1}{n_j-1}(\sum_{i=1}^{n} X_i^2 - \frac{1}{n_j}(\sum_{i=1}^{n} X_i)^2)$ (however if someone uses the maximum likelihood estimator $S_j^2 = \overline{x_j^2} - (\overline{x_j})^2$ this will be acceptable with 0.1 points deducted if no comment is made on this)

a) $\mathrm{Var}[X_1] = 211.0161$, $\mathrm{Var}[X_2] = 140.3557$,
$10.44638 \pm 2.676\sqrt{180.5352 \cdot (1/30 + 1/23)} = 10.44638 \pm 9.965075$. The second time period appears to give the faster access time as 0 is below the confidence interval.

b) We would need to observe 0 inside the confidence interval.

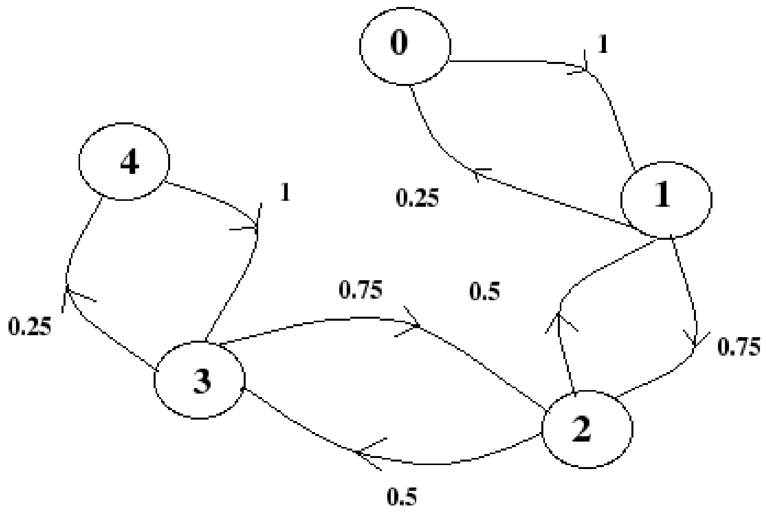
7. Lösning:

a) $62/200 = 0.31$, $76/190 = 0.4$ $0.31 - 0.4 = -0.09$, the estimator is unbiased as $\mathbf{E}[\frac{x}{n}] = \frac{1}{n}\mathbf{E}[x] = np/n = p$

b) $pn > 5$ or $(1-p)n > 5$ the probability of success cannot be too small and the sample needs to be at least 30 so we can use the normal approximation and plug–in $\hat{p}$ for p as then $p(1-p) \approx \hat{p}(1-\hat{p})$ in the confidence interval formula.

c) -0.09 ± 0.07969, yes I would be surprised as 0 is not in the confidence interval

8. Lösning:



a)

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0.25 & 0 & 0.75 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.75 & 0 & 0.25 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

b) An absorbing state is a state for which the probability of remaining in it is 1, a Markov chain is absorbing if it possess an absorbing state. If there is a one on $\mathbf{P}$'s diagonal then the Markov chain is absorbing.

c)

$$\mathbf{P}^2 = \begin{pmatrix} 0.25 & 0 & 0.75 & 0 & 0 \\ 0 & 0.625 & 0 & 0.375 & 0 \\ 0.125 & 0 & 0.75 & 0 & 0.125 \\ 0 & 0.375 & 0 & 0.625 & 0 \\ 0 & 0 & 0.75 & 0 & 0.25 \end{pmatrix}$$

$\mathbf{P}^n$ is the transition matrix of the change of the state of the Markov chain over n steps.

9. Lösning:

a) $\binom{n}{k}p^k(1-p)^{n-k}$, $k = 0, \ldots, n$, the binomial distribution is the sum of n independent trails of a Bernoulli distribution with parameter p

b) mean is np, variance $np(1-p)$

c) $\mathbf{E}[X] = \mu$ for every $\epsilon > 0$ $P(|X - \mu| \geq \epsilon) \leq \frac{\mathrm{Var}[X]}{\epsilon^2}$ $\mathbf{E}[X^2] < \infty$, proof see lecture.

d) multiply by n, use Chebyshev with mean and variance of binomial distribution. We use $\frac{S_n}{n}$ as an estimator for p, this gives us a rough bound on the probability that we will err in the estimation by at least ϵ