**Solutions to Exam April 2016**

**Part I**

1. No in general, only if the events are independent. This is part of Borel-Cantellis theorem.

2. Yes, this is part of theorem 10.1.1 about equivalences of weak convergences.

3. No not in general, if a chain is irreducible and aperiodic then there is a unique stationary measure.

4. Yes, Chebyshevs inequality is a generalization.

5. Yes, convergence almost surely implies convergence in probability which implies convergence in distribution.

**Part II**

**Exercise 1**

1. This is a Markov chain since the choice of choosing the next city does not depend on how many different cities were visited before. The state space is $S = \{1, ..., m\}$ and the transition probabilities

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} \frac{i}{m} & \text{if } j = i \\ \frac{m-i}{m} & \text{if } j = i + 1 \end{cases}$$

and 0 otherwise.

2. Once the chain goes to the next state it never goes back, hence the chain is not irreducible. The state $m$ is absorbing so recurrent and the other ones are transient. The period is 1 for all states, it is possible to reach each state in 1 step.

3.
$$\left(\frac{1}{2}, \frac{1}{2}, 0\right) \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} = \left(\frac{1}{18}, \frac{10}{18}, \frac{7}{18}\right)$$

4. $\tau_1 = 0$ for $m = 1$ and for $m = 2$ the $\tau_2$ is distributed according to a geometric distribution with parameter $\frac{1}{2}$. After visiting the first city the traveller has of probability of $\frac{1}{2}$ to visit the second city on the next day etc.

5. For $m > 1$ write $\tau_m$ as

$$\tau_m = \tau_1 + (\tau_2 - \tau_1) + ... + (\tau_m - \tau_{m-1})$$

$\tau_k - \tau_{k-1}$ is the waiting time until a never visited city has been visited provided that $k$ cities were visited before. As in (4) this is geometric random variable with parameter $\frac{m-k}{m}$, so $\mathbb{E}(\tau_m) = \sum_{k=1}^{m-1} \frac{m}{m-k}$.

6. We solve the system and get $\pi_k = 0$ for $k = 1, ..., m-1$ and $\pi_m = 1$ since $m$ is absorbing, we have that $\lim_{n\to\infty} \mathbb{P}(X_n = i) = \pi_i$.

**Exercise 2**

1. One possibility is let $f$ denote the density of the Gaussian variable

$$\mathbb{P}(X \geq x) = \int_x^\infty \frac{1}{t}(tf(t))dt = -\frac{1}{t}f(t)\Big|_x^\infty - \int_x^\infty \frac{1}{t^2}f(t)dt \leq \frac{1}{x}f(x)$$

and

$$\mathbb{P}(X \geq x) \geq \frac{1}{x}f(x) - \frac{1}{x^2}\int_x^\infty f(t)dt = \frac{1}{x}f(x) - \frac{1}{x^2}\mathbb{P}(X \geq x)$$

which yields the bound

2. Remark that we have $\mathbb{P}(S_n \geq xn) = (1+\epsilon_n)\frac{1}{\sqrt{2\pi n}}e^{-n\frac{x^2}{2}}$ as soon as

$$\frac{1}{n}\log\left(\sqrt{2\pi n}\cdot\mathbb{P}(S_n \geq xn)\right) = -\frac{x^2}{2} + o(n) \qquad (1)$$

and $o(n) := \frac{1}{n}\log(1+\epsilon_n)$ which goes to 0 as $n \to \infty$. From point 1 we have that

$$\mathbb{P}(S_n \geq nx) \leq \frac{1}{\sqrt{2\pi n}}e^{-n\left(\frac{x^2}{2}+\frac{\log(x)}{n}\right)} = \frac{1}{\sqrt{2\pi n}}e^{-n\left(\frac{x^2}{2}+o(n)\right)}$$

and

$$\mathbb{P}(S_n \geq nx) \geq \frac{1}{\sqrt{2\pi n}}e^{-n\left(\frac{x^2}{2}+\frac{1}{n}\log\left(\frac{x}{x^2+1}\right)\right)} = \frac{1}{\sqrt{2\pi n}}e^{-n\left(\frac{x^2}{2}+o(n)\right)}$$

and we can see that (1) is satisfied.

3. Use point 2 and calculate

$$\frac{1}{n}\log(\mathbb{P}(S_n \geq xn)) = \frac{1}{n}\left(\log(1+\epsilon_n) - \log(\sqrt{2\pi n}) - n\frac{x^2}{2}\right) \underset{n\to\infty}{\longrightarrow} -\frac{x^2}{2}$$

4. From a simple calculation we get $\Lambda(t) = \frac{t^2}{2}$. Call $g(t) = tx - \frac{t^2}{2}$, then the maxima are given by $x = t$ (take the derivative of $g$), then $\Lambda^*(x) = \frac{x^2}{2}$.

**Exercise 3**

1. For $x = 0$, we know that due to the CLT $\mathbb{P}(S_n \geq 0) \to \frac{1}{2}$ as $n \to \infty$, hence the lhs of equation (1) is going to 0, from the definition $I(0) = 0$. For $x > 1$, we have $\mathbb{P}(S_n > n) = 0$ so $\log(\mathbb{P}(S_n > 0)) = -\infty$. For $x = 1$ we have $\mathbb{P}(S_n = n) = 2^{-n}$ so the l.h.s. equals $-\log(2) = -I(1)$.

2. The distribution of $\frac{S_n+n}{2}$ is $Bin(\frac{1}{2}, n)$ on $\{0, ..., n\}$, hence

$$\mathbb{P}(S_n \geq nx) = \mathbb{P}\left(\frac{S_n+n}{2} \geq \frac{n(1+x)}{2}\right) = \sum_{k\geq n(1+x)/2}\frac{1}{2^n}\binom{n}{k}$$

3. For $k \geq \frac{n}{2}$, the map $k \mapsto \binom{n}{k}$ is monotonically decreasing.The lower bound is trivial since we are summing over $k$, for the upper bound consider that we have at most $n$ summands (even less!).

4.

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\log(Q_n(x)) &= \lim_{n\to\infty}\frac{1}{n}\log\left(\frac{n!}{a_n(x)!\cdot(n-a_n(x))!}\right)\\ &= \lim_{n\to\infty}\frac{1}{n}\log\left(\frac{n^n}{a_n(x)^{a_n(x)}(n-a_n(x))^{n-a_n(x)}}\right)\\ &= -I(x)+\log(2)\end{aligned}$$

5. Point 3 + 4 imply the claim of the proof. It means asymptotically that $\mathbb{P}(S_n \geq nx) \approx e^{-nI(x)}$.

6. The cumulant generating function is equal to $\Lambda(t) = \log(\cosh(t))$. We define $g(t) = tx - \log(\cosh t)$ and determine the maxima at $x = \tanh(t)$ or for $t = \tanh^{-1}(x) = \frac{1}{2}\log\left(\frac{1+x}{1-x}\right)$. Thus

$$\Lambda^*(x) = x\cdot\frac{1}{2}\log\left(\frac{1+x}{1-x}\right) - \log(\frac{1}{\sqrt{1-x}}) - \log(\frac{1}{\sqrt{1+x}})$$

simplifying yields the claim.

7. Let $Y_i$ denote a large claim. It appears with probability $\frac{1}{2}$.

$$\mathbb{P}\left(\sum_{i=1}^{30} Y_i \geq 15\right) \approx e^{-30\cdot I(\frac{1}{2})} \approx e^{-30\cdot 0.13} = 0.019.$$

**Exercise 4**

1. The law is equal to

$$\mu_X = \frac{1}{2}\mu_{0,\sigma^2} + \frac{1}{3}Exp(1/4) + \frac{1}{6}Geo\left(\frac{1}{3}\right)$$

The mean is equal to

$$\mathbb{E}(X) = \frac{1}{2}0 + \frac{1}{3}4 + \frac{1}{6}3 = \frac{11}{6}$$

and variance

$$\mathbb{E}(X^2) - \frac{11^2}{6^2} = \frac{1}{2}\sigma^2 + \frac{1}{3}2\cdot 16 + \frac{1}{6}15 = \frac{1}{2}\sigma^2 + \frac{79}{6} - \frac{11^2}{6^2} = \frac{1}{2}\sigma^2 + \frac{353}{36}$$

2. The CLT can be applied, so $\frac{1}{\sqrt{nVar(X)}}(S_n - n\frac{11}{6})$ converges to a standard normal variable.

3. If $\sigma = 2$, then $Var(X) = \frac{425}{36} = 11.80$.

$$\mathbb{P}\left(\frac{S_{100} - 100\frac{11}{6}}{10\sqrt{11.80}} > \frac{200 - 100\frac{11}{6}}{10\sqrt{11.80}}\right) \approx \mathbb{P}(Z > 0.48) = 0.31$$

4. We can use Berry-Esséen $\frac{3\mathbb{E}(|X_1 - \mathbb{E}(X_1)|^3)}{\sigma^3\sqrt{n}}$,

$$\mathbb{E}\left(\left|X_1 - \frac{11}{6}\right|^3\right) = \int_{11/6}^{\infty}\left(x - \frac{11}{6}\right)^3 \mu_X(dx) + \int_{-\infty}^{11/6}\left(\frac{11}{6} - x\right)^3 \mu_X(dx)$$

The first integral is equal to

$$\frac{1}{2}\int_{11/6}^{\infty}\left(x - \frac{11}{6}\right)^3 \mu_{(0,4)}(dx) + \frac{1}{3}\int_{11/6}^{\infty}\left(x - \frac{11}{6}\right)^3 \frac{1}{4}e^{-\frac{1}{4}x}dx + \frac{1}{6}\sum_{k\geq 2}\left(k - \frac{11}{6}\right)^3\left(\frac{2}{3}\right)^{k-1}\frac{1}{3}$$
$$= 0.44 + 80.93 + 8.79 = 90.16$$

and the second to

$$\frac{1}{2}\int_{-\infty}^{11/6}\left(\frac{11}{6} - x\right)^3 \mu_{(0,4)}(dx) + \frac{1}{3}\int_{0}^{11/6}\left(\frac{11}{6} - x\right)^3 \frac{1}{4}e^{-\frac{1}{4}x}dx + \frac{1}{6}\left(\frac{11}{6} - 1\right)^3\frac{1}{3}$$
$$= 14.52 + 0.21 + 0.03 = 14.76$$

hence the error is $\frac{3 \cdot 104.92}{2^3 \cdot 10} = 3.93$, big!

5. The law is $N(\sum_{i=1}^n i^{-2}, \sigma^2)$. We will show that $Y_n$ is a Cauchy-sequence, i.e. for all $\epsilon > 0$
$$\lim_{n\to\infty} \mathbb{P}(\sup_{m>n} |Y_n - Y_m| > \epsilon) = 0,$$
indeed since $Y_n$ is increasing,
$$\mathbb{P}(\sup_{m>n} |Y_n - Y_m| > \epsilon) \leq \frac{\mathbb{E}(\sup_{m>n} |Y_n - Y_m|)}{\epsilon} \leq \frac{\mathbb{E}(\sum_{i=n}^{\infty} U_{i^{-2}})}{\epsilon} = \frac{1}{\epsilon}\sum_{i=n}^{\infty} i^{-2}$$
which is going to 0 since the series is convergent.

6. We can compare the average energy consumption. For factory 1 it is $20 \cdot \frac{11}{6} = 36.66$ and for the factory 2 we have $\mathbb{E}(Y_{30}) = 0.56$ (you can also use bounds for $Y_{30}$), so much smaller consumption.