Sketch of solutions to Sheet 8

October 1, 2015

- Try not to consult these before you have tried the questions thoroughly.
- Very likely the solutions outlined below only represent a tiny subset of all possible ways of solving the problems. You are highly encouraged to explore alternative approaches!

1. Markov's inequality: For a positive random variable X and a constant $a > 0$, $\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}$.

   Observe $\mathbb{P}(X \geqslant a) = \mathbb{P}(X + b \geqslant a + b) = \mathbb{P}((X + b)^2 \geqslant (a + b)^2)$ (taking square on both side preserves the inequality since both side are non-negative). Apply Markov's inequality to the last expression and use the fact that $\mathbb{E}(X) = 0$ to obtain the result.

   Use differentiation or any other methods you like to find the minimum of $f(b) := \frac{\sigma^2+b^2}{(a+b)^2}$ to get the second inequality.

2. The waiting time X is a random variable with mean $\mu = 10$ and variance $\sigma^2 = 4$.

   Apply the inequality from question 1 to the random variable $X - 10$ (which has a zero mean and variance 4) and set $a = 2$, we get $\mathbb{P}(X \geqslant 12) = \mathbb{P}(X - 10 \geqslant 2) \leqslant \frac{4}{4+2^2} = 1/2$.

   If we just use Markov's inequality, then $\mathbb{P}(X \geqslant 12) \leqslant \frac{\mu}{12} = 5/6$. The inequality from question 1 gives a sharper bound (which makes sense because Markov's inequality does not make use of the information of variance!).

3. Observe that $\mathbb{P}(X \geqslant a) = \mathbb{P}(e^{tX} \geqslant e^{ta})$ and apply Markov's inequality to the last expression (we need $t \geqslant 0$ to preserve the inequality).

   Work out the mgf of $Poi(\lambda)$ and the RHS becomes $\exp(\lambda(e^t - 1) - ta)$. The expression in the exponent will be minimised at $t = \ln(a/\lambda) > 0$. Setting t as this value gives the last inequality.

4. Here $X \sim Bin(10000, 0.5)$. Here 10000 is "large" and we approximate X by a normal distribution $N(np, np(1 - p)) = N(5000, 2500)$:

   (a) $\mathbb{P}(X > 6000) \approx \mathbb{P}(Z > \frac{6000+0.5-5000}{\sqrt{2500}}) = \mathbb{P}(Z > 20.01) = 1 - \Phi(20.01) \approx 0$.

   (b) $\mathbb{P}(X \leqslant 5100) \approx \mathbb{P}(Z \leqslant \frac{5100+0.5-5000}{\sqrt{2500}}) = \mathbb{P}(Z \leqslant 2.01) = \Phi(2.01)$.

   (c) $\mathbb{P}(|X - 5000| < 50) = \mathbb{P}(4950 < X < 5050) \approx \mathbb{P}(\frac{4950+0.5-5000}{\sqrt{2500}} < Z < \frac{5050-0.5-5000}{\sqrt{2500}}) = \mathbb{P}(-0.99 < Z < 0.99) = \Phi(0.99) - \Phi(-0.99)$.

5. For a sequence of i.i.d random variables $X_1, X_2, ...$ with mean μ and finite variance σ^2, apply Chebyshev's inequality to $S_n := \frac{1}{n}\sum_{i=1}^{n} X_i$, and observe that $\mathbb{E}(S_n) = \mu$ and $\text{var}(S_n) = \frac{\sigma^2}{n}$. We obtain $\mathbb{P}(|S_n - \mu| \geqslant \epsilon) \leqslant \frac{\sigma^2}{n\epsilon^2}$. Taking limit gives the result.

6. (a) It follows by direct calculation by using the fact that $\mathbb{E}(X^3) = m^{(3)}(0)$.

(b) For fixed n, $skew(X)$ will tend to ∞ and $-\infty$ when $p \to 0$ and $p \to 1$ respectively. In these cases, the binomial distribution will be highly asymmetric, and thus a symmetric normal distribution might not be a good approximation. (Comment: Of course if n is really really big then $skew(X)$ can get closer to zero and we restore better symmetry of the binomial distribution. At the end, central limit theorem is just a limiting result of $n \to \infty$ and the theorem does not tell us how fast the convergence is. As hinted by this exercise, the converging performance may depend crucially on the interaction between p and n.)

7. Central limit theorem: for X_1, X_2, $X_3, \ldots$ a sequence of i.i.d. random variables with equal mean μ and finite variance σ^2, we have $\frac{\sum_{k=1}^n X_k/n-\mu}{\sigma/\sqrt{n}}$ converges to a standard normal random variable in distribution as $n \to \infty$, i.e

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\sum_{k=1}^n X_k/n - \mu}{\sigma/\sqrt{n}} \leqslant x\right) = \Phi(x)$$

where Φ is the CDF of a $N(0,1)$ random variable.

Let $X_1, X_2, \ldots$ be a sequence of i.i.d $Poi(1)$ random variables, such that $\mathbb{E}(X_i) = \text{var}(X_i) = 1$. Let $Y = \sum_{i=1}^n X_i$. Then Y has a $Poi(n)$ distribution. By CLT,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{Y/n - 1}{1/\sqrt{n}} \leqslant x\right) = \Phi(x)$$

In particular when we set $x = 0$, we have $1/2 = \Phi(0) = \lim_{n\to\infty} \mathbb{P}\left(\frac{Y/n-1}{1/\sqrt{n}} \leqslant 0\right) = \lim_{n\to\infty} \mathbb{P}(Y \leqslant n)$. But $Y \sim Poi(n)$, and thus $\mathbb{P}(Y \leqslant n) = \sum_{k=1}^n \mathbb{P}(Y = k) + \mathbb{P}(N = 0) = \sum_{k=1}^{\infty} \frac{e^{-n}n^k}{k!} + e^{-n}$. The result follows by taking limit on both side.