

## Exercise Sheet 7

### A: Warm-up questions

1.

$$G(z) = \sum_{r=1}^{\infty} z^r \{\pi(1-\pi)^{r-1}\} = \pi z \sum_{r=1}^{\infty} \{(1-\pi)z\}^{r-1} = \pi z \{1 - (1-\pi)z\}^{-1}.$$

$$G'(z) = \pi \{1 - (1-\pi)z\}^{-2} \text{ so } \mu = G'(1) = 1/\pi$$

$$G''(z) = 2\pi(1-\pi)\{1 - (1-\pi)z\}^{-3} \text{ so } G''(1) = 2(1-\pi)/\pi^2$$

$$\text{Therefore } \sigma^2 = G''(1) + G'(1) - \{G'(1)\}^2 = (1-\pi)/\pi^2.$$

2. The mgf of  $\log X$  is  $E(e^{s \log X}) = E(X^s) = \{\Gamma(\nu)\}^{-1} \lambda^\nu \int_0^\infty x^{s+\nu-1} e^{-\lambda x} dx$   
 $= \lambda^{-s} \Gamma(s+\nu)/\Gamma(\nu)$  [substitute  $u = \lambda x$ ].

### B: questions handed in

1. (a) This is a standard application of the result ... derived in lectures. ... to obtain that  $Y_1 \sim N(0, 1/2)$ .  
 (b) This mimicks the derivation of the above standard result closely:

$$\begin{aligned} M_Y(s) &\stackrel{\text{definition}}{=} \dots = \exp \left( (A^T s)^T \mu + \frac{1}{2} (A^T s)^T \Sigma (A^T s) \right) \\ &= \exp \left( s^T A \mu + \frac{1}{2} s^T A \Sigma A^T s \right) \end{aligned}$$

...

Finally, it remains to insert the numbers for  $A$  and  $\Sigma$  which yields  $Y \sim N \left( 0, \begin{pmatrix} 1/2 & 0 \\ 0 & 6 \end{pmatrix} \right)$

- (c) ... Therefore, the right solutions are

$$R \in \left\{ \begin{pmatrix} \sqrt{2} & -1/\sqrt{2} \\ 0 & \sqrt{3/2} \end{pmatrix}, \begin{pmatrix} -\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} \end{pmatrix}, \begin{pmatrix} \sqrt{2} & -1/\sqrt{2} \\ 0 & -\sqrt{3/2} \end{pmatrix}, \begin{pmatrix} -\sqrt{2} & 1/\sqrt{2} \\ 0 & -\sqrt{3/2} \end{pmatrix} \right\}$$

- (d) ...

$$\lambda \in \{1, 3\}$$

... The diagonalization of  $\Sigma$  is then

$$\Sigma = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

To show the required properties first note that ...

(e) My random numbers are

$$(x_1, x_2) = (1.22259, -0.597286)$$

..., so the sample is

$$(y_1, y_2) = \dots = \begin{pmatrix} 1.729003 \\ -1.596025 \end{pmatrix}$$

Similarly, setting  $Y = SX$  does the trick and the result is

$$(y_1, y_2) = \dots = \begin{pmatrix} 1.888711 \\ -1.263407 \end{pmatrix}$$

2. (a) i. Since  $S - 36$  is  $\text{Bin}(36, 0.5)$ , its mean and variance are 18 and 9 respectively. Hence  $\mathbb{E}[S] = 54$  and  $\text{Var}(S) = 9$ , so the normal approximation of the binomial delivers

$$P(S \in [49, 55]) = \dots \approx 0.5827$$

- ii. The continuity correction was introduced in STAT1005 (at the latest), it basically says that since  $P(S + c < 56) = P(S < 56)$  for all  $c \in [0, 1)$ , we should approximate the probability by selecting  $c$  in the middle, i.e.  $c = 1/2$ . This leads to

$$P(S \in [49, 55]) = P(48.5 < S \leq 55.5) \approx \dots \approx 0.6581$$

- iii. Without continuity correction, the result is not a very good approximation (not even the first digit is correct). With continuity correction, the result is accurate to three decimals.
- (b) This question shows that even the continuity correction cannot help when the CLT is grossly misunderstood.

i.

$$P(\bar{X} \leq 0) = \dots \approx \Phi\left(-\sqrt{n \frac{\delta}{1-\delta}}\right)$$

- ii. The continuity correction needs to interpolate between 0 and the next possible result, which is  $\alpha/n$ . Hence we obtain:

$$P(\bar{X} \leq 0) = \dots \approx \Phi\left(\frac{1/2n - \delta}{\sqrt{\delta(1-\delta)/n}}\right)$$

- iii. Setting  $\alpha = \sqrt{\frac{n^3}{n-1}}$  ensures that all  $X_i$  have unit variance. Also, the expectation tends to one as  $n \rightarrow \infty$ , so there are no singularities in the expectation or variance.

In the non-continuity corrected case we get

$$\lim_{n \rightarrow \infty} \Phi\left(-\sqrt{n \frac{\delta}{1-\delta}}\right) \dots = \Phi(-1) \approx 0.15866$$

In the continuity-corrected case we get

$$\lim_{n \rightarrow \infty} \Phi\left(\frac{1/2n - \delta}{\sqrt{\delta(1-\delta)/n}}\right) = \dots = \Phi(-1/2) \approx 0.3085$$

Comparing to the given true value of  $\lim_{n \rightarrow \infty} P(\bar{X} = 0) = 1/e \approx 0.3679$ , this is still about 6 percentage points off, in spite of the sample size having been sent to infinity.

The explanation of the problem lies in the fact that ...