Session 8: Solutions

Computer based exercise

Question 1

a)

A^2 has a skewed distribution.

	From sample*	From theoretical distribution of χ_1^2
mean	1.04	1
variance	2.07	2
50th percentile	0.47	0.45
90th percentile	2.79	2.71
95th percentile	3.92	3.84

* Seed was set arbitrarily to 16743. Percentiles of χ_1^2 were obtained using e.g. `di invchi2tail(1,0.05)` for the 95th percentile.

b) Distribution of E:

	From sample	From theoretical distribution of χ_3^2
mean	3.11	3
variance	6.62	6
50th percentile	2.44	2.37
90th percentile	6.65	6.25
95th percentile	8.33	7.81

c) Distribution of $A^2/(E/3)$: a $\frac{\chi_1^2/1}{\chi_3^2/3}$ variable should have an $F_{1,3}$ distribution (see lecture notes p 8.8).

	From sample	From theoretical distribution of $F_{1,3}$
50th percentile	0.60	0.59
90th percentile	5.06	5.54
95th percentile	8.94	10.1

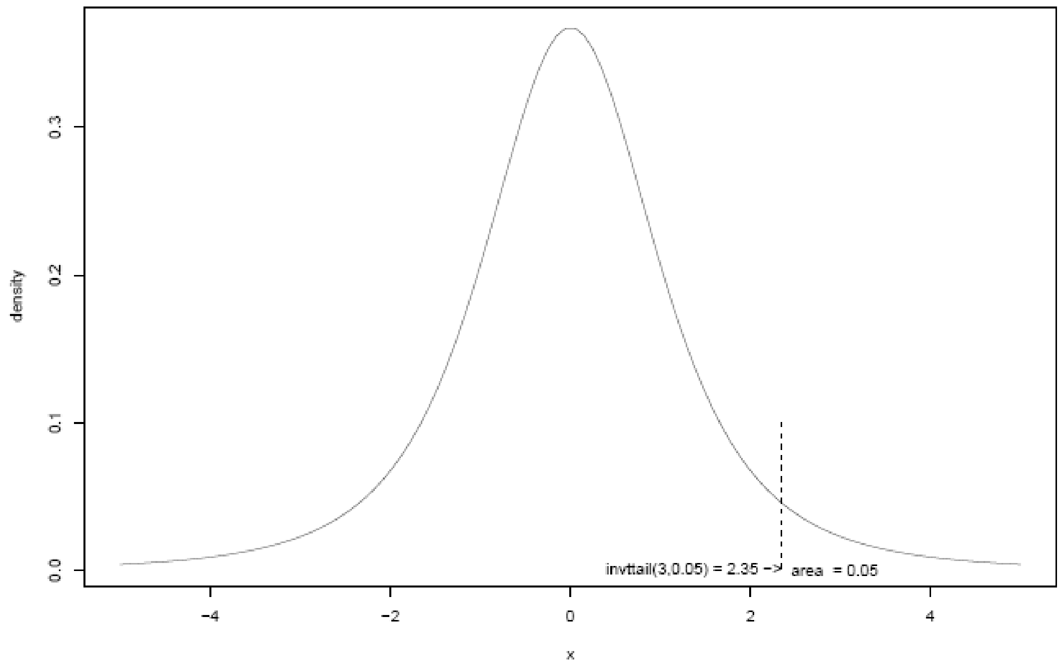
Percentiles of $F_{1,3}$ were obtained using e.g. `di invFtail(1, 3, 0.05)` for the 95th percentile.

The problem with using B^2 instead of A^2 is that since E is a function of B, these two random variables are not independent.

d) A random variable which is $+\sqrt{\frac{\chi_1^2/1}{\chi_3^2/3}}$ should have a $+\sqrt{F_{1,3}} = |t_3|$ distribution: note the positive square root.

	From sample	From theoretical distribution of t_3
50th percentile	0.78	0.76
90th percentile	2.25	2.35
95th percentile	2.99	3.18

Percentiles of $+\sqrt{F_{1,3}} = |t_3|$ can be obtained using e.g. `di invFtail(1, 3, 0.05)^0.5` or `di invttail(3, 0.025)` for the 95th percentile. Note that if using `invttail` it is necessary to look at the tail area above the 2.5th percentile to get the 95th percentile of the $|t_3|$ distribution (since both the lower and upper tail of the t_3 distribution will contribute to the upper tail of the $|t_3|$ distribution – see figure on next page).



The figure above shows the theoretical t_3 distribution. 2.35 is the theoretical 95th percentile of the t_3 distribution or the 90th percentile of the $|t_3|$ distribution.

Question 2

(a) /10 Seed set to 909693

	From sample	From theoretical distribution
mean	1.01	10/10 = 1
variance	0.21	20/100 = 0.2

(b) /50

	From sample	From theoretical distribution
mean	1.01	50/50 = 1
variance	0.04	100/2500 = 0.04

/100

	From sample	From theoretical distribution
mean	1.01	100/100 = 1
variance	0.02	200/10000 = 0.02

<u>More detail for a), b)</u>:

$$E\left(\frac{\chi_n^2}{n}\right) = \frac{n}{n} = 1; \qquad Var\left(\frac{\chi_n^2}{n}\right) = \frac{1}{n^2}Var(\chi_n^2) = \frac{1}{n^2}2n = \frac{2}{n}$$

hence the mean is 1 for any n; and the variance is $2/10$ for $n = 10$, $2/50$ for $n = 50$ and $2/100$ for $n = 100$.

(c) A χ_n^2/n distributed variable gets less skewed as n increases; the mean of the variable is 1 and the variance is $2/n$ which tends to 0 as n increases.

Thus:

$$t_n = \sqrt{\frac{\chi_1^2/1}{\chi_n^2/n}} \to \sqrt{\chi_1^2} \equiv N(0,1)$$

as n increases.

More detail for c): consider χ_n^2/n as n increases: since its variance tends to 0, the distribution of the random variable will tend to a 'spike' at its mean value, 1. So $\chi_n^2/n \to 1$ as n increases. Hence

$$t_n = \sqrt{\frac{\chi_1^2/1}{\chi_n^2/n}} \to \sqrt{\frac{\chi_1^2/1}{1}} = \sqrt{\chi_1^2} \equiv N(0,1)$$

as $n \to \infty$.