

Part I

2) The 90th, 95th and 99th percentiles of the Z-statistics: 1.726, 2.287 and 3.313 respectively

The 90th, 95th and 99th percentile of the standard normal distribution are: 1.281, 1.644 and 2.326 respectively

The 90th, 95th and 99th percentile of the t distribution 1.281, 1.644 and 2.326 respectively

4) For thresholds $T=2, 2.5, 3, 3.5$, calculate (a) the proportion of observed Z-scores that are greater than T in magnitude: The required values are 0.0725, 0.0368, 0.01706 and 0.00718 respectively

(b) the proportion of absolute Z-scores that you would expect to be greater than T in magnitude if all of the genes had equal means in the two groups (using an appropriate reference distribution): The required values are 0.0227, 0.0062, 0.00134, 0.00023 respectively

The FDR are 0.313, 0.168, 0.0791, 0.0323 respectively

Part II

- (i) Calculate the number of distinct IP addresses appearing within each minute, and calculate the 10th, 25th, 75th, and 90th percentiles of this value: 4233, 4402, 4843 and 39980 respectively
- (ii) Calculate the mean number of distinct times that each IP address appears within a minute of log data, and calculate the 10th, 25th, 75th, and 90th percentiles of these values: 0.00645, 0.01290, 0.01290 and 0.08387