Exercise Sheet 7

1. In an experiment on water quality, 16 independent water samples were taken downstream of a sewage processing plant, and X_i (in μgNH$_4$-N/l), the concentration of ammonia found in each, was measured. The average of these samples was found to be $\overline{x} = 204.2$.

We would subsequently like to know whether these data show (at level 5%) that the concentration of ammonia exceeds the legal limit, which is μgNH$_4$-N/l.

a) Assume that the standard deviation of such measurements is known from prior studies to be 10 μgNH$_4$-N/l.
Carry out a z test using this assumption, and ascertain whether or not a breach of the limit can be shown.
Write down the model assumptions, H_0 and H_A, the rejection set, the value of the test statistic and the outcome of the test explicitly.

b) How probable is this proof (at 5%) of the breach of the limit with 16 independent water samples if the true ammonia concentration is 205 μgNH$_4$-N/l?

c) How probable is mistakenly showing a breach of the limit using 16 independent water samples (at 5%) when the true ammonia concentration is exactly at the limit of 200 μgNH$_4$-N/l?

2. In this exercise, we shall investigate the effect of the Central Limit Theorem by means of simulation. We start out with a random variable X whose distribution is as follows: each of the values 0, 10 and 11 have probability $\frac{1}{3}$. Then we simulate the distribution of X and the distribution of the average of multiple copies of X.

a) Simulate X. Plot its distribution using a histogram of 1000 realizations of X and compare them to the normal distribution using the normal plot.

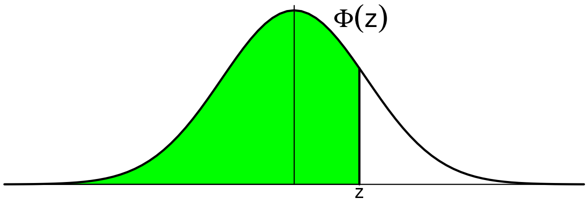
```
> par(mfrow=c(4,2)) # Several plots in a block
> werte <- c(0,10,11) # Possible values of X
> sim <- sample(werte, 1000, replace = T) # Simulate X
> hist(sim, main = "Original") # Create a histogram
> qqnorm(sim) # Create a normal plot
```

b) Now simulate $\overline{X} = \frac{X_1+X_2+X_3+X_4+X_5}{5}$, where the random variables X_i each have the same distribution of X_i and are independent. die gleiche Verteilung haben wie X und unabhängig sind. Plot the distribution of $\overline{X}$ using 1000 samples $\overline{X}$, and compare it to the normal distribution.

```
> n <- 5
> ## Simulate X_1, ..., X_n and put into a matrix with n rows (columns)
> sim <- matrix(sample(werte, n*1000, replace = T), ncol=n)
> sim.mean <- apply(sim, 1, "mean") # Compute row means
> hist(sim.mean, main = paste("Means of", n, "observations"))
> qqnorm(sim.mean)
```

c) Now simulate the distribution of $\overline{X}$ when $\overline{X}$ is the average of 10 or 200 copies of X_i, respectively.

Tabulated values of the cumulative normal distribution $\Phi(z) = \mathrm{P}\left[Z \leq z\right], \; Z \sim \mathcal{N}(0,1)$



Example: $\mathrm{P}\left[Z \leq 1.96\right] = 0.975$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767