Exercise 1

a) Validity is about whether the method in question is measuring what it is supposed to measure. Reliability (in the way we have discussed it) is about to what extent repeated use of the same method under the same conditions will give the same answer.

b) Questionnaire: There are two main problems related to validity in this method. First; the term "physical activity" is not clearly defined, so respondents may interpret it differently. Second; there is always a risk over over-reporting / under-reporting in such questionnaires. Reliability should be reasonably good.
Smartwatch: If we trust the watch, the validity of this method is good, as the mentioned output are objective measures. However, it is still unclear what measure (or combination of measures) to use. For reliability, the critical issue is that the subjects are asked to wear the watch for one day, while we are asking about habitual physical activity. One will expect rather large day-to-day variation, depending a bit on the target population.

c) We would prefer the method with the higher validity. A low validity will mean we are not able to estimate what we are aiming for, while a low reliability will mean that our estimate will have a larger uncertainty (which is considered less serious and can be compensated).

Exercise 2

a) Education is often used as a proxy for socioeconomic status, as it is a relatively simple measure. In Norway, you can get this information from registries. However, it is measuring only a part of socioeconomic status, so when used alone, it does not necessarily give a complete picture.

b) Both variables are categorical, so we would typically present them by frequencies and percentages. Bar charts may also be relevant here.

In good health	Freq.	Percent	Cum.
No	409	59.19	59.19
Yes	282	40.81	100.00
Total	691	100.00	

Education level	Freq.	Percent	Cum.
Secondary school	192	27.79	27.79
High school	312	45.15	72.94
College or university (1-4 år)	133	19.25	92.19
College or university (> 4 år)	54	7.81	100.00
Total	691	100.00	

c) We are to relate two categorical variables to each other, and our choice of method (as being taught in the course) is the chi-square test. This test gives a p-value of 0.002 (see output below), clearly less than the conventional significance level (0.05), and we conclude that there is a significant association between Edu and Health.

d)

	In good health		
Education level	No	Yes	Total
Secondary school	133 69.27	59 30.73	192 100.00
High school	181 58.01	131 41.99	312 100.00
College or university	70 52.63	63 47.37	133 100.00
College or university	25 46.30	29 53.70	54 100.00
Total	409 59.19	282 40.81	691 100.00

Pearson chi2(3) = 14.3413 Pr = 0.002

Looking at the within-row percentages, we observe an increasing trend in good health with increasing level of education (30.7%, 42.0%, 47.4%, 53.7%).

e) We have learned three different measures of association for a case like this; risk difference (RD), risk ratio (RR) and odds ratio (OR).

RD: 0.537 – 0.307 = 0.23, so an increase in 23 percentage points for those with the higher education as compared with those with the lower education.

RR: 0.537/0.307 = 1.75, so a 75% increased probability of good health for those with the higher education as compared with those with the lower education (or the probability is 1.75 times as high).

OR: $\frac{29/25}{59/133} = 2.61$, so the odds of good health is 2.6 times as high among those with the higher education as compared with those with the lower education.

Here, I have been comparing those with the higher education to those with the lower education. Turing it around would be equally relevant, giving measures RD = -0.23, RR = 0.57, OR = 0.38.

Instead of focusing on good health, we could focus on poor health:

RD: 133/192 – 25/54 = 0.693 – 0.463 = 0.23, thus a 23 percentage points higher probability for those with the lower education.

RR: 0.693/0.463 = 1.50, so 50% increased probability of poor health among those with lower education.

OR: $\frac{133/59}{25/29} = 2.61$, so the odds of poor health is 2.6 times as high among those with lower education.

Also here we could turn it around, comparing high with low education ….

Exercise 3

a) This is a box plot, summarizing the observations in the two groups. It gives median values (the horizontal lines within the boxes), interquartile range (the boxes) and minimum – maximum values, with the exception that Stata defines some observations as being outliers. We observe that it doesn't seem to be big differences between the genders.

b) There are two things to comment on. First; for males, the upper part of the box is close to BMI = 30. Above this box we have 25% of all the observations in the group, meaning close to 25% of the males will be considered obese. Second; the outliers. Some of them are extremely high, above 60 for two of them. It would make sense to check these to see if there is something wrong (and it isn't …).

c) It does not seem to be any association between age and BMI in these data. The estimated regression coefficients are both very close to zero (0.006 for males and -0.01 for females) and the p-values are far from being significant (0.83 and 0.58).

Exercise 4

a) The most striking observation here is that we see a clearly increasing trend among females over the years, while the trend among males is pretty flat. There was a difference between genders already among those born in 1870, with twice as many females reaching 100 years as males, and this difference has clearly increased among those born in 1900 and later.

b) We observe the same trend for both genders; normal weight seems stable, while overweigh is decreasing and obesity is increasing. This can be interpreted as more and more people moving from the overweight group into the obesity group. The main difference between genders is that more women belong to the normal weight group as compared to men.

Exercise 5

We have focused quite a bit on sampling variability as a source of variation during the course, and the confidence interval is set up to capture this type of variability (or uncertainty). There are obviously other types of uncertainty, e.g. related to quality of measurements and to model uncertainty (which is a bit beyond what we have discussed).