**Practical exercise using probability for continuous data following a normal distribution**

Please do as many questions as possible, make sure you are comfortable with the solutions, otherwise, contact the lecturer who covered this lecture.

Given the standard normal distribution, find:

1. The area under the curve between 
2. The probability that a picked at random will have a value between 
3. 
4. 
5. 
6. 
7. 
8. 
9. 
10. Given the following probabilities, find :
11.  (b) (c) 

(d) (e) 

11. Over a 25 year period the mean height of adult males increased from 175.8 cm to 179.1 cm, but the standard deviation stayed at 5.84 cm. The minimum height requirement for men to join the police force is 172 cm. What proportion of men would be too short to become policemen at the beginning and end of the 25 year period, assuming that the height of adult males has a normal distribution?

12. The Uptimer is a custom-made lightweight battery-operated activity monitor that records the amount of time an individual spends in the upright position. In a study of children ages 8 to 15 years, Eldridge et al., (2003)[[1]](#footnote-1) studied 529 normally developing children who each wore the Uptimer continuously for a 24-hour period that included a typical school day. The researchers found that the amount of time children spent in the upright position followed a normal distribution with mean of 5.4 hours and standard deviation of 1.3 hours. Assume that this finding applies to all children 8 to 15 years of age. Find the probability that a child selected at random spends less than 3 hours in the upright position in a 24-hour period.

13. If the total cholesterol values for a certain population are approximately normally distributed with a mean of 200 mg/100 ml and a standard deviation of 20 mg/100 ml, find the probability that an individual picked at random from this population will have a cholesterol value:

(a) Between 180 and 200 mg/100 ml (b) Greater than 225 mg/100 ml (c) less than 150 mg/100 ml (d) Between 190 and 210 mg/100 ml.

14. The following problem comes from the real study conducted to treat hypertension using antihypertensive drug therapy conducted by Castelli and Anderson, (1986)[[2]](#footnote-2). This was an Australian national study of risk factor prevalence. Two of the populations investigated are men whose blood pressures are within a normal (use of word “normal” here is not in the statistical sense) or accepted range and who are not taking any corrective medication, and men who have had high blood pressure but who are at present undergoing antihypertensive drug therapy.



Figure 1. Distribution of diastolic blood pressure for two populations of Australian males, Castelli and Anderson, (1986)

The population of men who are not taking corrective medication, diastolic blood pressure is approximately normally distributed with mean mm Hg and standard deviation mm Hg. For the men who are using antihypertensive drugs, diastolic blood pressure is approximately normally distributed with mean mm Hg and standard deviation mm Hg (see Figure 1 above).

In this assignment, interest is to be able to determine whether a man has normal blood pressure or whether he is taking antihypertensive medication solely on the basis of his diastolic blood pressure reading. This seemingly unimportant exercise is valuable in that it provides us with a foundation for hypothesis testing.

One thing we can see from Figure 1 above is that there is large amount of overlap between the two normal curves and as such, it will be difficult to distinguish between them.

1. Assume further that our first goal is to identify 90% of the individuals who are currently taking medication, what value of diastolic blood pressure should be designated as the lower cut off point? In other words, you must identify the value of diastolic blood pressure that marks off the lower 10% of this distribution (***hint: value marked x=? in figure 1)***. [15 marks]
2. For whatever value you find in (a), describe what the proportion below and above this cut off point represents. [15 marks]
3. Using the same cut-off point found in (a), what proportion of the men with normal blood pressures will be incorrectly labelled as antihypertensive drug users? [15 marks]
4. What would we call the error arising from your answer in part (c)? [10 marks]
5. By most people’s description, the error found in part (d) would be described as big. Therefore, in an effort to reduce this error, supposing you as the biostatistician analysing this data suggests to the investigators that we need the cut-off point for identifying individuals who are currently using antihypertensive drugs to be raised to 90 mm Hg. Using this new cut off point for the population with normal blood pressures, what proportion of men with normal blood pressures will be incorrectly labelled as antihypertensive drug users? [25 marks]
6. There is a penalty to pay in raising the cut-off point as suggested in part (e), estimate the proportion of men correctly labelled as using antihypertensive medication using the cut-off value suggested in part (e) and describe the consequences of your decision. [20 marks]

1. Eldridge B, Galena M, et al., (2003). Uptime Normative Values in Children Aged 8 to 15 Years. *Developmental Medicine and Child Neurology*, **45** : 189-193. [↑](#footnote-ref-1)
2. Catelli, W. P., and Anderson, K., “Antihypertensive Treatment and Plasma Lipoprotein Levels: The Associations in Data from a Population study”. *American Journal of Medicine Supplement,* Volume 80, February 14, 1986, 23-32. [↑](#footnote-ref-2)