Université Joseph Fourier – Grenoble I

L2 : Licence Sciences et Technologies

UE STA230 : midterm exam November 7$^{\text{th}}$, 2013

with answers

Duration 2h. Autorized documents: statistics tables and one A4 sheet handwriten (two pages). Calculators are authorized.

Exercise 1. (4pts) A group of high-altitude natives has been observed. Among them, 64% were Tibetans, the others were Aymara (from Bolivia). The proportion of persons having a low hemoglobin concentration (i.e. below 16 gm/dl), was 75% among Tibetans, 25% among Aymaras. For a person chosen at random in the sample, let A and B be the two events:

A: that person is Aymara,
B: that person has a low hemoglobin concentration.

1. Give the values of $\mathbb{P}[A]$, $\mathbb{P}[B \mid A]$ and $\mathbb{P}[B \mid \overline{A}]$. $\boxed{0.36, 0.25, 0.75}$

2. Compute $\mathbb{P}[B \text{ and } A]$, and $\mathbb{P}[B \text{ and } \overline{A}]$. $\boxed{0.09, 0.48}$

3. What proportion of the whole group had low hemoglobin concentration? $\boxed{0.57}$

4. Knowing that Mr. T. had a high hemoglobin concentration, what are the chances that he came from Bolivia? $\boxed{0.628}$

Exercise 2. (7pts) Actually, the hemoglobin concentration measurements of Mr. T. fluctuated independently from one day to another. It was estimated that the probability that he has a low hemoglobin concentration on a given day is 0.7.

1. Let X be the number of days Mr. T.'s hemoglobin concentration was low, out of 5 different days.

   1a) What probability distribution do you propose for X? Give its parameters. $\boxed{\mathcal{B}(5, 0.7)}$

   1b) Compute the probability that Mr. T's hemoglobin concentration is measured as low on at least 3 out of 5 consecutive days. $\boxed{0.8369}$

2. Mr. T.'s hemoglobin concentration has been measured on 120 days. Let Y be the random variable equal to the number of low measurements.

   2a) What probability distribution do you propose for Y? Give its parameters. $\boxed{\mathcal{B}(120, 0.7)}$

   2b) The probability distribution of Y is approximated by a normal distribution. What theoretical result justifies this approximation?

2c) Give the parameters of the approximating normal distribution. $\boxed{\mathcal{N}(84, 25.2)}$

2d) Using the normal approximation, compute the probability that at least 80 measurements were low. $\boxed{0.7872}$

2e) Using the normal approximation, find which number of days d is such that, with probability 0.90, Mr. T.'s measurements were low on at least d days. $\boxed{m \geqslant 78}$

Exercise 3. (9pts) The hemoglobin concentration in any given population follows a normal distribution. The results for means and standard deviations reported in the reference cited below are the following (unit: gm/dl).

- Tibetan men: $\mu = 15.6$, $\sigma = 1.5$

- Tibetan women: $\mu = 14.2$, $\sigma = 1.1$

- Aymara men: $\mu = 19.1$, $\sigma = 1.7$

- Aymara women: $\mu = 17.8$, $\sigma = 2.1$

1. What proportion of Tibetan men have a hemoglobin concentration lower than 16? $\boxed{0.6051}$

2. What proportion of Aymara women have a hemoglobin concentration lower than 16? $\boxed{0.1957}$

3. If a Tibetan man and an Aymara woman are taken at random, what are the chances the Tibetan man has a lower hemoglobin concentration than the Aymara woman? $\boxed{0.8030}$

4. What value of hemoglobin concentration is such that 90% of Tibetan women are below? $\boxed{15.61}$

5. What interval of hemoglobin concentration, centered on the expectation, contains 90% of Aymara men? $\boxed{[16.3\,;\,21.9]}$

6. The values given above were actually estimated on large enough samples. For Tibetan men the sample size was $n = 75$: $\overline{X} = 15.6$, $\sqrt{S^2} = 1.5$. Give a 95% confidence interval for the expected hemoglobin concentration among Tibetan men. $\boxed{[15.26\,;\,15.94]}$

7. All results in the reference below are given in terms of intervals $[\overline{X} - \frac{\sqrt{S^2}}{\sqrt{n}}\,;\,\overline{X} + \frac{\sqrt{S^2}}{\sqrt{n}}]$. What is the confidence level of such an interval? $\boxed{0.6827}$

8. On a sub-sample of 14 Aymara women, a mean hemoglobin concentration of 18.7, with standard deviation 1.6 was observed. Assuming that the theoretical σ is unknown, give a 98% confidence interval for μ. $\boxed{[17.52\,;\,19.88]}$

9. For that same sample, give a 98% confidence interval for σ. $\boxed{[1.13\,;\,2.96]}$

Reference: C. M. Beall et al. Hemoglobin concentration of high altitude Tibetans and Bolivian Aymara, American Journal of Physical Anthropology, 106:385–400 (1998)