Exam 2, Mathematics 2501, Section TR3 21 copies
3:40–4:50 pm, March 26, 2026, IH-1105
Instructor: Attila Máté[1]

1. Let p be a real with $0 < p < 1$ and let $q = 1 - p$. Let U be a discrete random variable such that $\mathrm{P}(U = 1) = p$ and $\mathrm{P}(U = 0) = q$. In all parts of this problem, you must show and explain your calculations. Giving only the final answers will get zero credit even if the answers are correct.

a) Find the expectation of U.

Solution. We have

$$\mathrm{E}(U) = 0 \cdot \mathrm{P}(U = 0) + 1 \cdot \mathrm{P}(U = 1) = 0 \cdot q + 1 \cdot p = p.$$

b) Find the variance of U.

Solution. We saw in part *a*) that $\mathrm{E}(U) = p$. As $U^2 = U$, we also have $\mathrm{E}(U^2) = \mathrm{E}(U) = p$. Hence

$$\mathrm{Var}(U) = E(U^2) - \big(E(U)\big)^2 = p - p^2 = p(1-p) = pq.$$

c) Using the result in part *a*), find the expectation of a binomial variable $\mathrm{Bin}(n, p)$.

Solution. Let X be variable with distribution $\mathrm{Bin}(n, p)$. Then X can be written as

$$X = \sum_{i=1}^{n} U_i,$$

where the variables U_i for i with $1 \le i \le n$ are independent variables with the same distribution as U above Hence

$$\mathrm{E}(X) = \sum_{i=1}^{n} \mathrm{E}(U_i) = \sum_{i=1}^{n} p = np.$$

d) Using the result in part *b*), find the variance of a binomial variable $\mathrm{Bin}(n, p)$.

Solution. Similarly as in part *c*, we have

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(U_i) = \sum_{i=1}^{n} pq = npq.$$

Solution.

2. A fair die is tossed 450 times. Let X denote the number of times that a number ≤ 4 is obtained.

a) Write a formula for the probability $\mathrm{P}(X = k)$ $(0 \le k \le 450)$.

Solution. X has a binomial distribution $X \sim \mathrm{Bin}(450, 2/3)$. That is,

$$P(X = k) = \binom{450}{k} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{450-k}.$$

b) Write the sum describing the probability $\mathrm{P}(X \le 305)$

[1]All computer processing for this manuscript was done under Debian Linux. $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX was used for typesetting. The Perl programming language was used in creating the $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX source file.

Solution. We have

$$P(X \le 305) = \sum_{k=0}^{305} P(X = k) = \sum_{k=0}^{305} \binom{450}{k} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{450-k}.$$

c) Find a numerical approximation for the probability $P(X \le 305)$; make sure to write the formula you are using to obtain the approximation.

Solution. We use normal approximation with continuity correction. A $\mathrm{Bin}(n, p)$ distribution has expectation np and standard deviation $\sqrt{np(1-p)}$. Accordingly, writing $D(X)$ for the standard deviation of X, we have

$$E(X) = 450 \cdot \frac{2}{3} = 300 \quad \text{and} \quad D(X) = \sqrt{450 \cdot \frac{2}{3} \cdot \frac{1}{3}} = 10.$$

Thus, the distribution of X can be approximated by a that of a normal variable $Y \sim \mathcal{N}(300, 10^2)$, i.e., by a normal variable Y having expectation 300 and standard deviation 10. Hence, using continuity correction,

$$P(X \le 305) \approx P(Y \le 305.5) = P\left(\frac{Y-300}{10} \le \frac{305.5-300}{10}\right) = P\left(\frac{Y-300}{10} \le .55\right) = \Phi(.55) \approx .7088;$$

here, the Φ denotes the distribution function of the standard normal distribution $\mathcal{N}(0, 1)$, and the last exact equality holds since $(Y - 300)/10$ has a standard normal distribution.

Calculating the sum part *b*) directly, with a short program written in the programming language Maxima, we get a more precise answer

$$P(X \le 305) \approx 0.7074816661881398.$$

The program is online at this website as `exam2_probl3_4.pdf`.

3. The following four measurements of the distance between two points are taken:

$$42.1, \qquad 45.3, \qquad 37.9, \qquad 41.5.$$

This is regarded as a random sample from a distribution $\mathcal{N}(m, \sigma^2)$, where m is the true distance, and σ^2 measures the precision of the method.

a) Find an unbiased estimate for m.

Solution. Given a random sample $x_1, x_2, \ldots, x_n$ (i.e., observations of the identically distributed independent random variables $X_1, X_2, \ldots, X_n$), an unbiased estimate for the mean is

$$\bar{x} \overset{def}{=} \frac{1}{n} \sum_{k=1}^{n} x_k.$$

In the present case, $n = 4$ and $x_1 = 42.1$, $x_2 = 45.3$, $x_3 = 37.9$, and $x_4 = 41.5$. Hence $\bar{x} = 41.7$ is an unbiased estimate for m.

b) Find an unbiased estimate for σ^2 if it is known that $m = 40.0$.

Solution. We have $\sigma^2 = \mathrm{Var}(X_k) = E\big((X_k - m)^2\big)$ (for all k with $1 \le k \le n$, since the variables X_k are identically distributed). Given that m is known, an unbiased estimate for σ^2 us the same as an unbiased estimate for the mean (expectation) $E\big((X_k - m)^2\big)$. That is, the estimate for σ^2 we are looking for is

$$\frac{1}{n} \sum_{k=1}^{n} (x_k - m)^2.$$

With the present data, $m = 40.0$ and x_k and n as above, we have approximately 9.7899 as the estimate for σ^2 we are looking for.

c) Find an unbiased estimate for σ^2 if m is unknown.

Solution. If m is not known, an unbiased estimate for σ^2 is

$$s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2,$$

where $\bar{x}$ is as above. In the present case, this gives $s^2 \approx 9.199$.

4. The following for measurements of the distance between two points are taken:

$$42.1, \qquad 45.3, \qquad 37.9, \qquad 41.5.$$

This is regarded as a random sample from a distribution $\mathcal{N}(m, \sigma^2)$, where m is the true distance, and σ^2 measures the precision of the method; the values of m and σ^2 are not known. Find a 95% confidence interval for m.

Solution. The level $1-\alpha$ confidence interval for the mean of a normal distribution given that the standard deviation is not known is

$$\left(\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right),$$

where n is the size of the sample, $\bar{x} = (1/n)\sum_{k=1}^{n} x_k$,

$$s = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(x_i - \bar{x})^2}\,;$$

further, given a random variable Y with a Student-t distribution with degrees of freedom f, and given y with $0 < y < 1$, the quantity $t_y(f)$ is defined such that

$$P\big(Y > t_y(f)\big) = y.$$

This confidence interval is discussed on p. 232 in the textbook.

In the present case, $\alpha = .05$, $n = 4$, $f = n - 1 = 3$, so $t_{\alpha/2}(n-1) = t_{.025}(3) = 3.18$, as given in the table on p. 326 in the textbook. Further, $x_1 = 42.1$, $x_2 = 45.3$, $x_3 = 37.9$, and $x_4 = 41.5$, and so $\bar{x} = 41.7$ and $s \approx 3.03315$. Hence, the 95% confidence interval for m in this case is

$$(36.87, 46.53).$$

5. The following for measurements of the distance between two points are taken:

$$42.1, \qquad 45.3, \qquad 37.9, \qquad 41.5.$$

This is regarded as a random sample from a distribution $\mathcal{N}(m, \sigma^2)$, where m is the true distance, and σ^2 measures the precision of the method; the values of m and σ^2 are not known.

a) Find a 95% confidence interval for σ^2.

Solution. The level $1-\alpha$ confidence interval for the variance of a normal distribution given that the mean and standard deviation is not known is

$$\left(\frac{1}{\chi^2_{\alpha/2}(n-1)}\sum_{k=1}^{n}(x_k - \bar{x})^2, \frac{1}{\chi^2_{1-\alpha/2}(n-1)}\sum_{k=1}^{n}(x_k - \bar{x})^2\right),$$

where n is the size of the sample and $\bar{x} = (1/n)\sum_{k=1}^{n} x_k$; further, given a random variable Y with a χ^2 distribution with degrees of freedom f, and given y with $0 < y < 1$, the quantity $t_y(f)$ is defined such that

$$P\big(Y > \chi_y^2(f)\big) = y.$$

This confidence interval is discussed on p. 235 in the textbook.

In the present case, $\alpha = .05$, $n = 4$, $f = n - 1 = 3$, so $\chi^2_{\alpha/2}(n-1) = \chi^2_{.025}(3) = 9.35$ and $\chi^2_{1-\alpha/2}(n-1) = \chi^2_{.975}(3) = .22$ as given in the table on p. 327 in the textbook. Further, $x_1 = 42.1$, $x_2 = 45.3$, $x_3 = 37.9$, and $x_4 = 41.5$, and so $\bar{x} = 41.7$ and $s \approx 3.03315$. Hence, the 95% confidence interval for σ^2 in this case is

$$(2.95, 128.0).$$

b) Find a 95% confidence interval for σ.

Solution. The confidence interval for σ is obtained by taking the square roots of the endpoints of the confidence interval for σ^2. In the present case, it is

$$(1.72, 11.31).$$