Solutions for Test Two, Math 477, Nov. 27, 2018

November 30, 2018

1. Twelve percent of the population is left-handed, so that, given any person selected randomly from the population, the probability that they are left handed is $p = 0.12$. Give the approximate probability that in a school of 200 students, there are at least 30 that are left handed. Discuss your assumptions.

SOLUTION Let X be the number of left handed students. Assume that the population of students does not have identical twins, or not that many, so that knowing whether student A is left or right handed tells us nothing about the whether student B is left or right handed. Then X is Bernoulli with $n = 200$ and success parameter $p = 0.12$. By the DeMoivre Laplace theorem,

$$Z := \frac{X - 200p}{\sqrt{200p(1-p)}}$$

is approximately standard normal.

Then

$$\begin{aligned} P(X \geq 30) &= P\left(X - 200p \geq 30 - 200p)\right) \\ &= P\left(\frac{X - 200p}{\sqrt{200p(1-p)}} \geq \frac{30 - 200p}{\sqrt{200p(1-p)}}\right) \\ &= P\left(Z \geq \frac{30 - 200p}{\sqrt{200p(1-p)}}\right) \ . \end{aligned}$$

Inserting $p = 0.12$ into the formula, we are left with evaluating $P(Z \geq 1.30558241...)$.

The full credit answer then is

$$P(X \geq 30) \approx 1 - \Phi(1.30558241...) \approx \int_{1.3056}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x \ .$$

2. The expected number of typographical errors per page of a magazine is 0.2. Compute the probability that the next page you read has:

(a) 0 typos

(b) 2 or more typos

SOLUTION The typos occur independently in the words on the page, so the number of typos per page can be expected to be binomial, then since the probability that a given word will contain a typo is small, we may make the Poisson approximation, and assume the that number of typos has a Poisson distribution with parameter λ. If X is Poisson with parameter λ, then $\mathrm{E}(X) = \lambda$, and hence we have that X, the number of typos per page is Poisson with parameter 0.2. It follows that

$$P(X = n) = e^{-1/5}\frac{5^{-n}}{n!} \ .$$

We can now easily answer the question:

(a) $P(X = 0) = e^{-1/5}$. That is, $P(X = 0) = .818730753...$

(b) $P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{6}{5}e^{-1/5}$. That is, $P(X \geq 2) = 0.017523096...$

3. An urn contains 4 white balls and 4 black balls. At each trial, 4 balls are drawn at random from the urn, and then replaced. The trial is a "success" if exactly two white balls and 2 black balls are drawn. Let N be the number of the trial on which the first success occurs. Compute

$$P(N = 5)\ , \qquad \mathrm{E}(N) \quad \text{and} \quad \mathrm{Var}(N)\ .$$

SOLUTION We may think of each drawing as a Bernoulli trial, with success probability p being the probability that the drawing yields exactly 2 white balls and 2 black balls. To compute p, note that there are $\binom{8}{4}$ ways to form a subset of 4 balls, corresponding the the 4 balls that are drawn. There are $\binom{4}{2}$ ways to form subsets consisting of 2 white balls, and then $\binom{4}{2}$ way to choose a subset of 2 black balls.

Hence

$$p = \frac{\binom{4}{2}^2}{\binom{8}{4}} = \frac{18}{35} = .514285714...\ .$$

Since the trials are independent, N is geometric with success parameter p. Then

$$P(N = 5) = (1-p)^4 p = \frac{1503378}{52521875} = 0.0286238448...,$$

and we know that

$$\mathrm{E}(N) = \sum_{n=1}^{\infty} n(1-p)^{n-1}p = \frac{1}{p} = \frac{35}{18}$$

and

$$\mathrm{Var}(N) = \sum_{n=1}^{\infty} n^2(1-p)^{n-1}p - (\mathrm{E}(N))^2 = \frac{1-p}{p^2} = \frac{595}{324}\ .$$

4. Let X_1, X_2, X_3, X_4 and X_5 be independent, and let each be exponentially distributed with parameter λ. For $a > 0$, compute

$$P(\min\{X_1, X_2, X_3, X_4, X_5\} \leq a) \qquad \text{and} \qquad P(\max\{X_1, X_2, X_3, X_4, X_5\} \leq a))\ .$$

Also, compute $\mathrm{E}(\min\{X_1, X_2, X_3, X_4, X_5\})$.

SOLUTION The probability density function for each X_j is $\lambda e^{-\lambda x}$ for $x > 0$, and 0 otherwise, and the cumulative distribution function is $F(x) = 1 - e^{-\lambda x}$. From the theory of order statistics, the probability density function for $Y := \min\{X_1, X_2, X_3, X_4, X_5\}$ is

$$\frac{5!}{4!}f(x)[1 - F(x)]^4 = -\frac{\mathrm{d}}{\mathrm{d}x}[1 - F(x)]^5\ ,$$

and so, for $a > 0$,

$$P(Y < a) = \int_0^a -\frac{\mathrm{d}}{\mathrm{d}x}[1 - F(x)]^5 \mathrm{d}x = 1 - [1 - F(a)]^5 = 1 - e^{-5\lambda a} \ .$$

We also compute, integrating by parts,

$$\mathrm{E}(Y) = \int_0^\infty x5f(x)[1 - F(x)]^4 \mathrm{d}x = \int_0^\infty e^{-5\lambda x} \mathrm{d}x = \frac{1}{5\lambda} \ .$$

Likewise, from the theory of order statistics, the probability density function for $Z := \max\{X_1, X_2, X_3, X_4, X_5\}$ is

$$\frac{5!}{4!}[F(x)]^4 f(x) = \frac{\mathrm{d}}{\mathrm{d}x}[F(x)]^5 \ ,$$

and so, for $a > 0$.

$$P(Z < a) = \int_0^a \frac{\mathrm{d}}{\mathrm{d}x}[F(x)]^5 \mathrm{d}x = [F(x)]^5 = (1 - e^{-\lambda a})^5 \ .$$

Alternatively, this one can be answered *without* using the theory of order statistics since

$$P(Z < a) = P(\cap_{j=1}^5 \{X_j < a\}) = \prod_{j=1}^5 P(X_j < a) = (1 - e^{-\lambda a})^5 \ .$$

5. Let X and Y be independent standard normal random variables.

(a) Define $U = X$ and $V = X/Y$. Compute the joint density of U and V, and find then the density of V.

(b) Compute $P(X \leq Y \leq 2X)$.

SOLUTION

a The joint probability density of X and Y is

$$f(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2} \ .$$

We have $u(x, y) = x$ and $v(x, y) = x/y$, so that

$$x(u, v) = u \qquad \text{and} \qquad y(u, v) = \frac{u}{v} \ .$$

We then compute the Jacobian

$$\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \frac{|u|}{v^2} \ .$$

Therefore, the joint probability density function of U and V is

$$g(u, v) = \frac{1}{2\pi} e^{-u^2(1+v^{-2})/2} \frac{|u|}{v^2} \ .$$

The probability density function for V, $g_V(v)$, then is

$$\begin{aligned} g_V(v) &= \int_{\mathbb{R}} g(u,v)\mathrm{d}u \\ &= \frac{1}{\pi v^2}\int_0^\infty e^{-u^2(1+v^{-2})/2} u\mathrm{d}u \\ &= \frac{1}{\pi v^2}\int_0^\infty e^{-t(1+v^{-2})}\mathrm{d}t \\ &= \frac{1}{\pi}\frac{1}{1+v^2} . \end{aligned}$$

This was all you needed to say for part **a**, but the computation means that V has a Cauchy distribution, and shows how the Cauchy distribution arises as the distribution of the ratio of two standard normal variables.

b We compute

$$P(X \le Y \le 2X) = \frac{1}{2\pi}\int_{x<y<2x} e^{-(x^2+y^2)/2}\mathrm{d}x\mathrm{d}y .$$

In terms of polar coordinates, $y = x$ corresponds to $\theta = \pi/4$ and $y = 2x$ corresponds to $\theta = \arctan(2)$. So

$$P(X \le Y \le 2X) = \frac{1}{2\pi}\int_{\pi/4}^{\arctan(2)}\left(\int_0^\infty e^{-r^2/2}\mathrm{d}r\right)\mathrm{d}\theta = \frac{\arctan(2) - \pi/4}{2\pi} .$$