Math 3070 § 1.   **Second Midterm Exam**   Name:____Solutions____
Treibergs *στ*   October 23, 2013

(1.) *In a 1996 study for the Journal of Statistics Education, it was found that the mean body temperature of humans is* $98.249°$ *and the standard deviation is* $0.729°$. *Assume that body temperatures are normally distributed. What is the probability that the temperature of a random human is less than* $98.600°$, *the usual standard used by physicians? (Give answer to at least three digits.) Below what temperature will 33% of the temperature samples be?*

Let X denote the temperature of a random human, which is given as a normally distributed random variable with mean $\mu = 98.249$ degrees and standard deviation $\sigma = 0.729$ degrees. By standardizing we find

$$P(X < 98.600) = P\left(Z = \frac{X - \mu}{\sigma} < \frac{98.600 - 98.249}{.729}\right) = \Phi(0.4814815)$$

Interpolating the Standard Normal Curve Areas Table A.3, we see that

$$0.4815 = .15(.4900) + .85(.4800)$$

so that

$$\Phi(.4815) \approx (.15)\Phi(.4900) + (.85)\Phi(.4800) = (.15)(.6879) + (.85)(.6844) = \boxed{0.6849}.$$

The second part asks for the 33% percentile η. Since the standard quantile satisfies $.3300 = \Phi(z_{.67}) = \Phi(-.4400)$ from Table A.3, we have $z_{.67} = -.4400$. Thus 33% of the temperatures will be below

$$\eta = \mu + z_{.67}\,\sigma = 98.249 + (-.4400)(0.729) = \boxed{97.928°}.$$

(2.) *Let X be a random variable whose cumulative distribution function (CDF) is*

$$F(x) = \begin{cases} 0, & \text{if } x < 2; \\ 2 - \dfrac{4}{x}, & \text{if } 2 \le x \le 4; \\ 1, & \text{if } 4 < x. \end{cases}$$

Find the probability that $3 \le X \le 5$. Find the median $\tilde{\mu}$. Find the formula for the probability density function $f(x)$. Find the mean and variance of X.

Since X is a continuous random variable,

$$P(3 \le X \le 5) = P(X \le 5) - P(X \le 3) = F(5) - F(3) = 1 - \left(2 - \frac{4}{3}\right) = \boxed{\frac{1}{3}}.$$

The median satisfies $\dfrac{1}{2} = F(\tilde{\mu}) = 2 - \dfrac{4}{\tilde{\mu}}$ so $\tilde{\mu} = \boxed{\dfrac{8}{3}}$.

The probability density is given by the Fundamental Theorem of Calculus as

$$f(x) = F'(x) = \boxed{\begin{cases} 0, & \text{if } x < 2; \\ \dfrac{4}{x^2}, & \text{if } 2 \le x \le 4; \\ 0, & \text{if } 4 < x. \end{cases}}$$

1

The mean and expected square are

$$E(X) = \int_{-\infty}^{\infty} x \, f(x) \, dx = \int_{2}^{4} x \cdot \frac{4}{x^2} \, dx = 4\big(\ln(4) - \ln(2)\big) = \boxed{4\ln(2)};$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \, f(x) \, dx = \int_{2}^{4} x^2 \cdot \frac{4}{x^2} \, dx = 4(4-2) = 8.$$

The variance is by the shortcut formula $V(X) = E(X^2) - E(X)^2 = \boxed{8 - 16\,(\ln(2))^2} = 0.313.$

(3.) *The lifetime of a GPS device manufactured by Caineville Cairns has an exponential distri-*
bution with a mean of 2 years. Find the probability that the GPS device lasts longer than three
years. Given that one of the GPS devices is already four years old but is still functioning, what
is the probability that it will cease functioning within the next year?

The lifetime is distributed as an exponential variable $X \sim \mathrm{Exp}(\lambda)$. Its mean is $E(X) = \frac{1}{\lambda} = 2$
so $\lambda = \frac{1}{2}$. Its cumulative density is $F(x) = 1 - e^{-\lambda x}$ for $x \geq 0$. The probability that the device
lasts longer than three years is

$$P(X > 3) = 1 - P(X \leq 3) = 1 - F(3) = 1 - \left(1 - e^{-\frac{1}{2}\cdot 3}\right) = \boxed{e^{-\frac{3}{2}}} = 0.223.$$

The conditional probability that it lasts less than five years given that it has lasted four years is

$$P(X < 5 \mid x \geq 4) = \frac{P(\{X < 5\} \text{ and } \{x \geq 4\})}{P(\{x \geq 4\})} = \frac{P(\{4 \leq X < 5\})}{1 - P(\{x < 4\})} = \frac{F(5) - F(4)}{1 - F(4)}$$

$$= \frac{\left(1 - e^{-5/2}\right) - \left(1 - e^{-4/2}\right)}{1 - \left(1 - e^{-4/2}\right)} = \frac{-e^{-5/2} + e^{-2}}{e^{-2}} = \boxed{1 - e^{-1/2}} = .393.$$

(4.) *The top recreational activity of Utahns is still the relaxing sport of fishing, with 41% of those*
surveyed indicating that they fished during the year. Suppose that 180 randomly selected Utahns
are asked whether they fished in the past year. What is the probability that fewer than 60 had
fished? Determine an expression for the exact probability. (You do not need to evaluate.) Using
an approximation, compute the probability. Why was your approximation valid?

Let X denote the number of people who fished last year among the $n = 180$ interviewed.
Assuming independence of trials, X is a binomial variable where the probability of success (had
fished) is $p = 0.41$. The probability that fewer than 60 have fished is

$$P(X < 60) = P(X \leq 59) = B(59, 180, 0.41) = \sum_{y=0}^{59} \binom{180}{y} (0.41)^y \, (0.59)^{180-y}.$$

We approximate by a normal distribution with continuity correction. Thus for $x = 59$,

$$P(X \leq x) \approx \Phi\left(\frac{x + \frac{1}{2} - np}{\sqrt{npq}}\right) = \Phi\left(\frac{59.5 - 180(0.41)}{\sqrt{180(0.41)(0.59)}}\right) = \Phi(-2.167114)$$

Interpolating in Table A.3, since $-2.1671 = (.71)(-2.17) + (.29)(-2.16)$ we have $P(X \leq 59) \approx$
$\Phi(-2.1671) \approx (.71)\Phi(-2.17) + (.29)\Phi(-2.16) = (.71)(.0150) + (.29)(.0154) = \boxed{0.0151}$.
The approximation is valid according to the rule of thumb in Devore, since both $nq > np = (180)(.41) = 73.8 \geq 10$.

(5.) *An urn contains seven red balls and three blue balls. Suppose that four balls are randomly chosen from the urn without replacement. Let X denote the number of red balls chosen. Find a formula for the probability mass function $p(x)$. What are $E(X)$ and $V(X)$? Let $h(x) = (-1)^x$. Find $E(h(X))$. Find the formula for cumulative mass function $F(x)$.*

The random variable is hypergeometric $X \sim$ Hypergeometric$(n = 4, M = 7, N = 10)$, that is $n = 4$ are chosen without replacement from $N = 10$ of which $M = 7$ are successes (red). Thus the probability mass function is

$$p(x) = \begin{cases} \dfrac{\binom{7}{x}\binom{3}{4-x}}{\binom{10}{4}}, & \text{if } x \in \{1, 2, 3, 4\}; \\ \\ 0, & \text{otherwise.} \end{cases}$$

Computing,

$$p(1) = \frac{\binom{7}{1}\binom{3}{3}}{\binom{10}{4}} = \frac{7 \cdot 1}{210} = \frac{1}{30}, \qquad p(2) = \frac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}} = \frac{21 \cdot 3}{210} = \frac{9}{30},$$

$$p(3) = \frac{\binom{7}{3}\binom{3}{1}}{\binom{10}{4}} = \frac{35 \cdot 3}{210} = \frac{15}{30}, \qquad p(4) = \frac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \frac{35 \cdot 1}{210} = \frac{5}{30}.$$

The fraction of successes is $p = M/N = .7$. The mean and variance of the hypergeometric variable is according to the formulas

$$E(X) = np = (4)(.7) = \boxed{2.8}. \qquad V(X) = npq\left(\frac{N-n}{N-1}\right) = (4)(.7)(.3)\left(\frac{10-4}{10-1}\right) = \boxed{.56}$$

The expectation of $h(X)$ is $E(h(X)) =$

$$= \sum_{x=1}^{4} h(x)p(x) = (-1)^1 p(1) + (-1)^2 p(2) + (-1)^3 p(3) + (-1)^4 p(4) = \frac{-1 + 9 - 15 + 5}{30} = \boxed{-\frac{1}{15}}.$$

The formula for cumulative mass function is $F(x) = \sum_{y \leq x} p(y)$ so

$$F(x) = \begin{cases} 0, & \text{if } x < 1; \\ \\ \dfrac{1}{30}, & \text{if } 1 \leq x < 2; \\ \\ \dfrac{10}{30}, & \text{if } 2 \leq x < 3; \\ \\ \dfrac{25}{30}, & \text{if } 3 \leq x < 4; \\ \\ 1, & \text{if } 4 \leq x. \end{cases}$$