Homework #8, Mei Maddox.

Mei Maddox

Submit the solution on Canvas into the corresponding assignment (e.g. "Homework #1") in the form of R Markdown report, knitted into either of the available formats (HTML, pdf or Word). Provide only code and output. NO NEED TO COPY THE PROBLEM FORMULATION (!)

Problem 1

6.28 Mean birth weight for boys is 3.41kg with a standard deviation of 0.55kg. Assume that the distribution of birth weight is approximately normal.

a) 0.0490084 proportion of baby boys are born with weight less than 2.5kg

b) The z-score of a baby boy that weighs 1.5kg is $z = \frac{1.5-3.41}{0.55} \approx$ -3.4727273.

c) The probability of a baby boy birth weight between 2.5 and 4.0 is 0.80929488212755, 8.98497811849308e-15, 1, OK, integrate(f = function(x) dnorm(x, 3.41, 0.55), lower = 2.5, upper = 4).

d) A baby weighing 3.6 at birth is in the 64th percentile.

e) A baby which falls at the 96th percentile weights approximately 4.3728773kg.

6.42 The probability of making any given free throw is 0.90. During a game, a player shot 10 free throws. Let X=number of free throws made.

a) In order to assume X has a binomial distribution, the probability of making any given free throw is the same across all throws, the outcome of each throw is independent of each other, there are a fixed number of throws, and the sample space for each throw is either making it (success) or missing it (failure).

b) For this example $n = 10$ and $p = 0.90$.

c)

- $P(X = 10) = \binom{10}{10} \times 0.90^{10} \times 0.1^0 \approx 0.3486784$
- $P(X = 9) = \binom{10}{9} \times 0.90^9 \times 0.1^1 \approx 0.3874205$
- $P(X > 7) \approx 0.9298092$

6.43 Refer to the previous exercise. Over the course of a season, this player shoots 400 free throws.

a) The mean of a binomial distribution for large number of trials (throws) is $\mu = np = 400 \times 0.90 = 360$. The standard deviation of a binomial distribution for large number of trials (throws) is $\sigma = \sqrt{400 \times 0.90 \times 0.1} = \sqrt{36} = 6$.

b) By the normal distribution approximation, the expected number of made throws would certainly fall between $360 - 3 \times 6 = 342$ and $360 + 3 \times 6 = 378$ because approx. 99.7% of the data falls within 3 standard deviations of the mean.

c) The proportion of made throws would likely fall between $342/400 = 0.855$ and $378/400 = 0.945$.

6.51

For following random variables, explain why at least one condition needed to use the binomial distribution is unlikely to be satisfied.

a) X = number people in a family of size 4 who go to church on a given Sunday, when any one of them goes 50% of the time in the long run. Family church-going decisions can be influenced by social groups and/or neighborhood location. Therefore, independence between each family cannot be assumed.

b) X = number voting for the Dem candidate out of the 100 votes in the precinct, when 60% of the population voted for the Dem in that state. Voting probabilities are not necessarily identical across all precincts in the same state. It therefore violates the rule that all probabilities must be the same across all trials.

c) X = number of females in a random sample of 4 students from a class of size 20, when half the class is female. Due to sampling variety, the number of females is not necessarily going to be the same across multiple trials.

Problem 2

1

a) $E[X] = \int_{-\infty}^{\infty} xp(x)dx$
$E[aX+b] = \int_{-\infty}^{\infty}(ax+b)p(ax+b)dx$
$E[aX+b] = \int_{-\infty}^{\infty} ax \times p(ax+b)dx + b \times p(ax+b)dx$
$E[aX+b] = \int_{-\infty}^{\infty} ax \times p(ax+b)dx + \int_{-\infty}^{\infty} b \times p(ax+b)dx$
$E[aX+b] = a\int_{-\infty}^{\infty} x \times p(ax+b)dx + b\int_{-\infty}^{\infty} p(ax+b)dx$
$E[aX+b] = a\int_{-\infty}^{\infty} x \times p(ax+b)dx + b \times 1$
$E[aX+b] = a\int_{-\infty}^{\infty} x \times p(x)dx + b \times 1$
$E[aX+b] = a\int_{-\infty}^{\infty} x \times p(x)dx + b$
$E[aX+b] = aE[X] + b$

b) Using the results from (a), show that for $Z = \frac{X-\mu}{\sigma}, X \sim N(\mu, \sigma^2)$ we have $E[Z] = 0$ and $Var[Z] = 0$. Keep in mind μ, σ are constants, and X - random variable, represent Z as $aX + b$, where a, b only contain constants. aE[X]+b

$E[Z] = E[\frac{X-\mu}{\sigma}]$
$E[Z] = E[\frac{X}{\sigma} + \frac{-\mu}{\sigma}]$
$= E[\frac{1}{\sigma}X + \frac{-\mu}{\sigma}]$
$= \frac{1}{\sigma}E[X] + \frac{-\mu}{\sigma}$
$= \frac{1}{\sigma}\mu + \frac{-\mu}{\sigma}$
$= \frac{\mu}{\sigma} + \frac{-\mu}{\sigma}$
$= 0$

$Var[Z] = Var[\frac{X-\mu}{\sigma}]$
$Var[Z] = Var[\frac{X}{\sigma} + \frac{-\mu}{\sigma}]$
$= Var[\frac{1}{\sigma}X + \frac{-\mu}{\sigma}]$
$= (\frac{1}{\sigma})^2 \times Var[X]$
$= \frac{1}{\sigma^2} \times Var[X]$
$= \frac{1}{\sigma^2} \times \sigma^2$

$= \frac{\sigma^2}{\sigma^2}$
$= 1$

2 For a Bernoulli random variable X

a) $E[X] = 1 \times p + 0 \times (1-p) = p + 0 = p$

b) $V[X] = \sum_i (x_i - E[X])^2 P(X = x_i)$
$= (1-p)^2 \times p + (0-p)^2 \times (1-p)$
$= (1 - 2p + p^2) \times p + p^2(1-p)$
$= p - 2p^2 + p^3 + p^2 - p^3$
$= p - p^2$
$= p(1-p)$

Problem 3

```
z.norm <- function(z.score, mu=0, sigma=1){
  a <- mu-z.score*sigma
  b <- mu+z.score*sigma
  integrate(function(x) dnorm(x, mu, sigma), a, b)
}
```

a) Probability of X landing within 1.64 standard deviations of the mean is approximately 0.898994833051792, 6.78610064346018e-14, 1, OK, integrate(f = function(x) dnorm(x, mu, sigma), lower = a, upper = b)

b) Probability of X landing within 2.58 standard deviations of the mean is approximately 0.990119968484459, 1.87606566781763e-08, 1, OK, integrate(f = function(x) dnorm(x, mu, sigma), lower = a, upper = b).

c) Probability of X landing within 0.67 standard deviations of the mean is approximately 0.49714220980938, 5.51938727843404e-15, 1, OK, integrate(f = function(x) dnorm(x, mu, sigma), lower = a, upper = b).

The probabilities of (a)-(c) would not change if the values of mu and sigma change because the z-score measurement is scaled in such a way that a change of amount z is the same, regardless of the original raw value range.

```
z.norm(2.58)
```

```
## 0.99012 with absolute error < 1.9e-08
```

```
z.norm(2.58, mu = 12)
```

```
## 0.99012 with absolute error < 1.9e-08
```

```
z.norm(2.58, sigma = 7)
```

```
## 0.99012 with absolute error < 1.9e-08
```

```r
z.norm(2.58, 12, 7)
```

```
## 0.99012 with absolute error < 1.9e-08
```