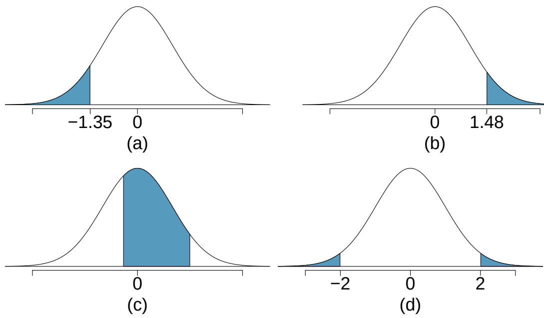
Practice Exercises: Lesson 2.3 Solutions

STAT 1201
Introduction to Probability and Statistics

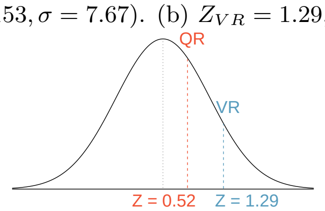
ONLINE AND DISTANCE EDUCATION

4 Distributions of random variables

4.1 (a) 8.85%. (b) 6.94%. (c) 58.86%. (d) 4.56%.



4.3 (a) Verbal: $N(\mu = 151, \sigma = 7)$, Quant: $N(\mu = 153, \sigma = 7.67)$. (b) $Z_{VR} = 1.29$, $Z_{QR} = 0.52$.



(c) She scored 1.29 standard deviations above the mean on the Verbal Reasoning section and 0.52 standard deviations above the mean on the Quantitative Reasoning section. (d) She did better on the Verbal Reasoning section since her Z-score on that section was higher. (e) $Perc_{VR} = 0.9007 \approx 90\%$, $Perc_{QR} = 0.6990 \approx 70\%$. (f) $100\% - 90\% = 10\%$ did better than her on VR, and $100\% - 70\% = 30\%$ did better than her on QR. (g) We cannot compare the raw scores since they are on different scales. Comparing her percentile scores is more appropriate when comparing her performance to others. (h) Answer to part (b) would not change as Z-scores can be calculated for distributions that are not normal. However, we could not answer parts (d)-(f) since we cannot use the normal probability table to calculate probabilities and percentiles without a normal model.

4.5 (a) $Z = 0.84$, which corresponds to approximately 159 on QR. (b) $Z = -0.52$, which corresponds to approximately 147 on VR.

4.7 (a) $Z = 1.2$, $P(Z > 1.2) = 0.1151$.
(b) $Z = -1.28 \rightarrow 70.6°$F or colder.

4.9 (a) $N(25, 2.78)$. (b) $Z = 1.08$, $P(Z > 1.08) = 0.1401$. (c) The answers are very close because only the units were changed. (The only reason why they differ at all because 28°C is 82.4°F, not precisely 83°F.) (d) Since $IQR = Q3 - Q1$, we first need to find $Q3$ and $Q1$ and take the difference between the two. Remember that $Q3$ is the 75^{th} and $Q1$ is the 25^{th} percentile of a distribution. Q1 = 23.13, Q3 = 26.86, IQR = 26. 86 - 23.13 = 3.73.

4.11 (a) No. The cards are not independent. For example, if the first card is an ace of clubs, that implies the second card cannot be an ace of clubs. Additionally, there are many possible categories, which would need to be simplified. (b) No. There are six events under consideration. The Bernoulli distribution allows for only two events or categories. Note that rolling a die could be a Bernoulli trial if we simply to two events, e.g. rolling a 6 and not rolling a 6, though specifying such details would be necessary.

4.13 (a) $0.875^2 \times 0.125 = 0.096$. (b) $\mu = 8$, $\sigma = 7.48$.

4.15 If p is the probability of a success, then the mean of a Bernoulli random variable X is given by
$\mu = E[X] = P(X = 0) \times 0 + P(X = 1) \times 1$
$= (1 - p) \times 0 + p \times 1 = 0 + p = p$

4.17 (a) Binomial conditions are met: (1) Independent trials: In a random sample, whether or not one 18-20 year old has consumed alcohol does not depend on whether or not another one has. (2) Fixed number of trials: $n = 10$. (3) Only two outcomes at each trial: Consumed or did not consume alcohol. (4) Probability of a success is the same for each trial: $p = 0.697$. (b) 0.203. (c) 0.203. (d) 0.167. (e) 0.997.

4.19 (a) $\mu 34.85$, $\sigma = 3.25$ (b) $Z = \frac{45 - 34.85}{3.25} = 3.12$. 45 is more than 3 standard deviations away from the mean, we can assume that it is an unusual observation. Therefore yes, we would be surprised. (c) Using the normal approximation, 0.0009. With 0.5 correction, 0.0015.

4.21 (a) $1 - 0.75^3 = 0.5781$. (b) 0.1406. (c) 0.4219. (d) $1 - 0.25^3 = 0.9844$.

4.23 (a) Geometric distribution: 0.109. (b) Binomial: 0.219. (c) Binomial: 0.137. (d) $1 - 0.875^6 = 0.551$. (e) Geometric: 0.084. (f) Using a binomial distribution with $n = 6$ and $p = 0.75$, we see that $\mu = 4.5$, $\sigma = 1.06$, and $Z = 2.36$. Since this is not within 2 SD, it may be considered unusual.

4.25 (a) $\overset{Anna}{1/5} \times \overset{Ben}{1/4} \times \overset{Carl}{1/3} \times \overset{Damian}{1/2} \times \overset{Eddy}{1/1} = 1/5! = 1/120$. (b) Since the probabilities must add to 1, there must be $5! = 120$ possible orderings. (c) $8! = 40{,}320$.

4.27 (a) 0.0804. (b) 0.0322. (c) 0.0193.

4.29 (a) Negative binomial with $n = 4$ and $p = 0.55$, where a success is defined here as a female student. The negative binomial setting is appropriate since the last trial is fixed but the order of the first 3 trials is unknown. (b) 0.1838. (c) $\binom{3}{1} = 3$. (d) In the binomial model there are no restrictions on the outcome of the last trial. In the negative binomial model the last trial is fixed. Therefore we are interested in the number of ways of orderings of the other $k - 1$ successes in the first $n - 1$ trials.

4.31 (a) Poisson with $\lambda = 75$. (b) $\mu = \lambda = 75$, $\sigma = \sqrt{\lambda} = 8.66$. (c) $Z = -1.73$. Since 60 is within 2 standard deviations of the mean, it would not generally be considered unusual. Note that we often use this rule of thumb even when the normal model does not apply. (d) Using Poisson with $\lambda = 75$: 0.0402.

4.33 (a) $\frac{\lambda^k \times e^{-\lambda}}{k!} = \frac{6.5^5 \times e^{-6.5}}{5!} = 0.1454$
(b) The probability will come to $0.0015 + 0.0098 + 0.0318 = 0.0431$ (0.0430 if no rounding error).
(c) The number of people per car is $11.7/6.5 = 1.8$, meaning people are coming in small clusters. That is, if one person arrives, there's a chance that they brought one or more other people in their vehicle. This means individuals (the people) are not independent, even if the car arrivals are independent, and this breaks a core assumption for the Poisson distribution. That is, the number of people visiting between 2pm and 3pm would not follow a Poisson distribution.

4.35 0 wins (-\$3): 0.1458. 1 win (-\$1): 0.3936. 2 wins (+\$1): 0.3543. 3 wins (+\$3): 0.1063.

4.37 Want to find the probability that there will be 1,786 or more enrollees. Using the normal approximation, with $\mu = np = 2,500 \times 0.7 = 1750$ and $\sigma = \sqrt{np(1-p)} = \sqrt{2,500 \times 0.7 \times 0.3} \approx 23$, $Z = 1.61$, and $P(Z > 1.61) = 0.0537$. With a 0.5 correction: 0.0559.

4.39 (a) $Z = 0.67$. (b) $\mu = \$1650$, $x = \$1800$. (c) $0.67 = \frac{1800-1650}{\sigma} \rightarrow \sigma = \$223.88$.

4.41 (a) $(1-0.471)^2 \times 0.471 = 0.1318$. (b) $0.471^3 = 0.1045$. (c) $\mu = 1/0.471 = 2.12$, $\sigma = \sqrt{2.38} = 1.54$. (d) $\mu = 1/0.30 = 3.33$, $\sigma = 2.79$. (e) When p is smaller, the event is rarer, meaning the expected number of trials before a success and the standard deviation of the waiting time are higher.

4.43 $Z = 1.56$, $P(Z > 1.56) = 0.0594$, i.e. 6%.

4.45 (a) $Z = 0.73$, $P(Z > 0.73) = 0.2327$. (b) If you are bidding on only one auction and set a low maximum bid price, someone will probably outbid you. If you set a high maximum bid price, you may win the auction but pay more than is necessary. If bidding on more than one auction, and you set your maximum bid price very low, you probably won't win any of the auctions. However, if the maximum bid price is even modestly high, you are likely to win multiple auctions. (c) An answer roughly equal to the 10th percentile would be reasonable. Regrettably, no percentile cutoff point guarantees beyond any possible event that you win at least one auction. However, you may pick a higher percentile if you want to be more sure of winning an auction. (d) Answers will vary a little but should correspond to the answer in part (c). We use the 10^{th} percentile: $Z = -1.28 \rightarrow \$69.80$.

4.47 (a) $Z = 3.5$, upper tail is 0.0002. (More precise value: 0.000233, but we'll use 0.0002 for the calculations here.)
(b) $0.0002 \times 2000 = 0.4$. We would expect about 0.4 10 year olds who are 76 inches or taller to show up.
(c) $\binom{2000}{0}(0.0002)^0(1-0.0002)^{2000} = 0.67029$.
(d) $\frac{0.4^0 \times e^{-0.4}}{0!} = \frac{1 \times e^{-0.4}}{1} = 0.67032$.

5 Foundations for inference

5.1 (a) Mean. Each student reports a numerical value: a number of hours. (b) Mean. Each student reports a number, which is a percentage, and we can average over these percentages. (c) Proportion. Each student reports Yes or No, so this is a categorical variable and we use a proportion. (d) Mean. Each student reports a number, which is a percentage like in part (b). (e) Proportion. Each student reports whether or not s/he expects to get a job, so this is a categorical variable and we use a proportion.

5.3 (a) The sample is from all computer chips manufactured at the factory during the week of production. We might be tempted to generalize the population to represent all weeks, but we should exercise caution here since the rate of defects may change over time. (b) The fraction of computer chips manufactured at the factory during the week of production that had defects. (c) Estimate the parameter using the data: $\hat{p} = \frac{27}{212} = 0.127$. (d) *Standard error* (or SE). (e) Compute the SE using $\hat{p} = 0.127$ in place of p: $SE \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.127(1-0.127)}{212}} = 0.023$. (f) The standard error is the standard deviation of $\hat{p}$. A value of 0.10 would be about one standard error away from the observed value, which would not represent a very uncommon deviation. (Usually beyond about 2 standard errors is a good rule of thumb.) The engineer should not be surprised. (g) Recomputed standard error using $p = 0.1$: $SE = \sqrt{\frac{0.1(1-0.1)}{212}} = 0.021$. This value isn't very different, which is typical when the standard error is computed using relatively similar proportions (and even sometimes when those proportions are quite different!).

5.5 (a) Sampling distribution. (b) If the population proportion is in the 5-30% range, the success-failure condition would be satisfied and the sampling distribution would be symmetric. (c) We use the formula for the standard error: $SE = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.08(1-0.08)}{800}} = 0.0096$. (d) Standard error. (e) The distribution will tend to be more variable when we have fewer observations per sample.