1 One

For a), the only possibility where $X+Y=0$ occurs when $X=0$ and $Y=0$ since X is binomial and Y bernoulli. Hence, $P(X+Y=0)=P(X=0)P(Y=0)$, and since $P(X=0)=(0.7)^9$ and $P(Y=0)=0.7$, we get $P(X+Y=0)=(0.7)^10$.

For b), we can calculate the pmf of Z as, for $a \in \{1, ..., 9\}$:

$$p_Z(a) = \sum_{i=0}^{1} p_X(a-i)p_Y(i) = p_X(a)p_Y(0) + p_X(a-1)p_Y(1) = \binom{9}{a}(0.3)^a(0.7)^{9-a}(0.7) + \binom{9}{a-1}(0.3)^{a-1}(0.7)^{10-}$$

$$p_Z(a) = (\binom{9}{a} + \binom{9}{a-1})(0.3)^a(0.7)^{10-a}$$

Since $\binom{9}{a} + \binom{9}{a-1} = \binom{10}{a}$ by the binomial coefficient's recursion, we get $p_Z(a) = \binom{10}{a}(0.3)^a(0.7)^{10-a}$. For the case of $a=0$, we see that $p_X(a-1) = p_X(-1) = 0$, so $p_Z(0) = p_X(0)p_Y(0) = (0.7)^10 = \binom{10}{0}(0.3)^0(0.7)^10$ as we found before.

On the other hand, for the case of $a=10$, $p_X(a) = p_X(10) = 0$, so $p_Z(10) = p_X(10-1)p_Y(1) = p_X(9)p_Y(1) = (0.3)^9(0.3) = \binom{10}{10}(0.3)^10(0.7)^0$. Since $X+Y$ cannot equal any other possible values given their own range of values, the pmf of Z is:

$$p_Z(z) = \begin{cases} \binom{10}{z}(0.3)^z(0.7)^{10-z}, z \in \{0, ..., 10\} \\ 0 \text{ otherwise} \end{cases}$$

But this is exactly the binomial distribution with $n=10$ and $p=0.3$, so $Z \sim \text{Bin}(10, 0.3)$. For c), observe that Z is the sum of the number of independent successful binary events described by the 9 of X, with the number of independent successful binary events described by the single one of Y.

Hence, without reference to b), we can conclude Z is the number of successful independent binary events out of 10 total events of the same probability $p=0.3$, and so must be distributed binomially.

2 Two

For a), note that there are only three possible arrangements of (X_1, X_2, X_3)– either the red ball is taken out first, second or third, so only one of the $x_i = 1$ while the other two are 0. Hence, as the other two balls are necessarily white, these describe all possible events, giving us a count of 3 events total.

We assume uniformness without more information, so the pmf vanishes everywhere but at these points, where it takes a value of 1/3:

$$p(x_1, x_2, x_3) = \begin{cases} 1/3, (x_1, x_2, x_3) \in \{(1,0,0),(0,1,0),(0,0,1)\} \\ 0 \text{ otherwise} \end{cases}$$

For b), the marginal distributions come from summing over all possible probabilities that give us $X_i = x$. For example, $X_1 = 0$ occurs when either $X_2 = 1$ and $X_3 = 0$ or $X_2 = 0$ and $X_3 = 1$, so $p_{X_1}(0) = p(0,1,0) + p(0,0,1) = 1/3 + 1/3 = 2/3$, while $X_1 = 1$ is only possible when both $X_2 = X_3 = 0$, so $p_{X_1}(0) = p(1,0,0) = 1/3$.

By symmetry of the three random variables X_i, the other two also have the same distribution, so we get the pmfs:

$$p_{X_1}(x) = p_{X_2}(x) = p_{X_3}(x) = \begin{cases} 1/3, x = 1 \\ 2/3, x = 0 \\ 0, \text{ otherwise} \end{cases}$$

which is just a Bernoulli distribution with $p = 1/3$, so $X_1, X_2, X_3 \sim \text{Ber}(1/3)$. For c), observe that $p(1,0,0) = 1/3$, while $p_{X_1}(1)p_{X_2}(0)p_{X_3}(0) = 1/3(2/3)^2$, so $p(1,0,0) \neq p_{X_1}(1)p_{X_2}(0)p_{X_3}(0)$, so the joint distribution does not equal the separated one, implying that they are not independent.

For d), observe that $Y = X_1 + X_2$ can only take on two values of 0 and 1, as only one of X_1 and X_2 can be 1 at a time while the other has to be 0. The case of $Y = 1$ occurs when either $X_1 = 1$ or $X_2 = 1$, so $p_Y(1) = p(1,0,0) + p(0,1,0) = 1/3+1/3 = 2/3$, while $Y = 0$ occurs when both are 0, so $p_Y(0) = p(0,0,1) = 1/3$ only.

Hence, as before, we see that $p_Y(y) = p_{X_1}(y)$, so it is also described by a Bernoulli distribution $Y \sim \text{Ber}(1/3)$. For e), observe we can use the same argument with $X_1 + X_2$ to get the pmf of $X_2 + X_3$, just swapping X_1 and X_3, since the distribution of (X_1, X_2, X_3) is symmetric under permutation. Hence, the pmf of $X_2 + X_3$ is the same as that of Y in d), so $X_2 + X_3 \sim \text{Ber}(1/3)$.

3 Three

For a), observe that the pmf of $D_T = D_1 + D_2$ can be calculated using the sum formula, using the fact that D_1 and D_2 both take integer values from 0 to ∞, for $a \in \mathbb{Z}_{\geq 0}$:

$$p_{D_T}(a) = \sum_{d_1+d_2=a} p_{D_1}(d_1)p_{D_2}(d_2) = \sum_{i=0}^{\infty} p_{D_1}(i)p_{D_2}(a-i)$$

since $p_{D_2}(a-i) = 0$ when $i > a$, we can truncate the sum to:

$$p_{D_T}(a) = \sum_{i=0}^{a} p_{D_1}(i)p_{D_2}(a-i) = \sum_{i=0}^{a} \frac{15^i e^{-15}}{i!} \frac{3^{a-i}e^{-3}}{(a-i)!}$$

Multiplying the numerator and denominator by $a!$ and pulling out $e^{-15}e^{-3}/a!$ gives us:

$$p_{D_{T(a)}} = e^{-18}/a! \sum_{i=0}^{a} \frac{a!}{i!(a-i)!} 15^i 3^{a-i}$$

and as the sum is the binomial expansion of 18^a, we get $p_{D_T}(a) = 18^a e^{-18}/a!$ for non-negative integer a, and 0 otherwise. This means that D_T is exactly distributed as a Poisson with rate $r = 18$, so $D_T \sim \text{Pois}(18)$.

For b), we are given $D_T = 25$, and we want to find the probability that $D_2 = d$, where $d = 4$ in this case. For c), we do this in general to find the pmf for $D_2 | D_T = 25$, which we can calculate as:

$$f(d) = P(D_2 = d | D_T = 25) = \frac{P(D_2 = d \cup D_T = 25)}{P(D_T = 25)}$$

Since $D_2 = d$ and $D_T = 25$ is equivalent to saying $D_2 = d$ and $D_1 = 25 - d$, and D_1 and D_2 are independent, the numerator probability is just $p_{D_1}(25-d)p_{D_2}(d)$, while the denominator is $p_{D_T}(25)$, so we get:

$$f(d) = \frac{(e^{-15}15^{25-d}/(25-d)!)(e^{-3}3^d/d!)}{e^{-18}18^{25}/25!}$$

$$f(d) = \frac{25!}{d!(25-d)!} \frac{15^{25-d}3^d}{18^25}$$

recognizing the binomial coefficient, and distributing the 18's over each power of 15 and 3, we get:

$$f(d) = \binom{25}{d}(3/18)^d(15/18)^{25-d} = \binom{25}{d}(1/6)^d(5/6)^{25-d}$$

for $d \in \{0, ..., 25\}$ and 0 otherwise, since D_2 cannot be larger than D_T and cannot be negative. This means that $D_2 | D_T = 25$ is binomially distributed with $n = 25$ and $p = 1/6$, so $D_2|D_T = 25 \sim \text{Bin}(25, 1/6)$. In fact, in genral, $D_2|D_T = a \sim \text{Bin}(a, 1/6)$, which we can derive simply by replacing 25 by a in our previous derivation, likewise $D_1|D_T = a \sim \text{Bin}(a, 5/6)$, where 1/6 represents the probability of a customer going to the second shop and 5/6 the first shop, given they went there.

For d), we simply are modelling the expectation of $D_1|D_T = 27$, which is $\text{Bin}(27, 5/6)$, and has an expected value of $E[D_1|D_T] = np = 27 \cdot 5/6 = 22.5$, so about 22.5 people will go to the first shop.

4 Four

For a), we compute as below, using integration by parts:

$$A = \int_0^\infty xe^{-x/10}dx = -10xe^{-x/10}|_0^\infty + 10\int_0^\infty e^{-x/10}dx = -100e^{-x/10}|_0^\infty = 100$$

For b), we observe that A is the expected value of $10X$, where $X \sim \text{Exp}(1/10)$, hence we expect that this distribution would be useful in estimating A. For c), in fact if we choose n samples $X_1, ..., X_n$ from Exp(1/10), a good estimator for A would be their average multiplied by 10:

$$T_n = \frac{10}{n}\sum_{i=1}^n X_i$$

since A is the expected value of $10X$ when X is exponentially distributed. We see that $E[T_n] = \frac{1}{n}\sum_{i=1}^n E[10X_i] = \frac{1}{n}(nA) = A$, so our desired property is satisfied.

Since the X_i are independent by assumption, we can compute the variance of T_n in d) as the separate summed variances of each $10X_i/n$, or:

$$\text{Var}(T_n) = \sum_{i=1}^n \text{Var}(10X_i/n)$$

since the covariances vanish. Using the scaling property of variances, we can pull out the constant, leaving us with:

$$\text{Var}(T_n) = \frac{100}{n^2}\sum_{i=1}^n \text{Var}(X_i)$$

as each X_i is exponentially distributed with $\lambda = 1/10$, the variance of X_i is $1/\lambda^2 = 100$, so:

$$\text{Var}(T_n) = \frac{100}{n^2}\sum_{i=1}^n 100 = \frac{100^2}{n}$$

For e), by Chebyshev's inequality, we see that $P(|T_n - A| > 0.5) \leq \frac{\text{Var}(T_n)}{0.5^2} = 4(100^2)/n$, so $P(|T_n - A| > 0.5) \leq 40000/n$. Choosing n such that $40000/n \leq 0.1$ gives us our desired inequality, so we just choose $n \geq 400000$. At the very minimum, we hence need a sample size of $n = 400000$ for T_n to estimate A within well within a range of 0.5 of its value with 90% confidence.

For f), we complete the code as follows:

```
N<- 40 0000
sims <-10 * rexp(N,1/10)
sum(sims)/N
```

For g), note that this estimate has a probability of not satisfying $|T_n - A| \leq 0.5$ of 0.1, so we would expect 10% of the time that this estimate would not be within a 0.5 interval of A, or that about 10 of the 100 times this experiment is repeated, it would not be satisfied.

5 Five

Code for a):

```
A <- 100
N <- 10000
m <- 100
sims <- numeric(m)
for (i in 1:m) {
    sims[i] <- sum(10 * rexp(N ,1/10))/N
}
p <- sum(abs(sims-A)>0.5)/m
p
bound <- 40000/N
bound

df <- data.frame(estimate=sims)
ggplot(df, mapping=aes(x=estimate, fill = abs(sims - A) > .5))+geom_histogram(binwidth = .25
```

We see that the bound is around 4, while the probability is 0.63 approximately, so it is definitely bounded above.Code for b):

```
N <- 1000
X1 <- numeric(N)
X2 <- numeric(N)
player <- function(player_cards, dealer_card) {
    score <- score_blackjack(player_cards)
    hasace <- 'A' %in% player_cards$numbers
    deal <- sum(dealer_card$value)
    decision <- TRUE
    if (score == 12 & ((2 <= deal & deal <= 3) | deal >= 7)) {
        decision <- TRUE
    }
    if (score == 12 & (4 <= deal & deal <= 6)) {
        decision <- hasace
    }
    if ((13<= score & score <= 16) & deal <= 6) {
        decision <- hasace
    }
    if ((13<= score & score <= 16) & deal >= 7) {
        decision <- TRUE
    }
    if (score==17) {
        decision <- hasace
    }
    if (score==18 & deal <= 8) {
        decision <- FALSE
```

```
    }
    if (score >= 19) {
        decision <- FALSE
    }

  return(decision)
}
play_a_round <- function(shuffled, position = 1) {
  # takes shuffled deck and the current position
  # of the top of the undealt deck as input arguments
  # deal first pairs
  # feel free to modify the function as needed
  player_hand <- c(position, position + 1)
  dealer_hand <- c(position + 2:3)
  position <- position + 4
  while (player(shuffled[player_hand, ], shuffled[dealer_hand[1], ])) {
    player_hand <- c(player_hand, position)
    position <- position + 1
  }
  player_score <- score_blackjack(shuffled[player_hand, ])
  if (player_score == 21) {
    return(c(TRUE, position)) # player wins immediately
  }
  if (player_score > 21) {
    return(c(FALSE, position)) # player loses immediately
  }
  while (dealer(shuffled[dealer_hand, ])) {
    dealer_hand <- c(dealer_hand, position)
    position <- position + 1
  }
  dealer_score <- score_blackjack(shuffled[dealer_hand, ])
  # return c(winner, current position)
  return(c(winner(player_score, dealer_score), position))
}
for (i in seq_len(N)) {
    shuffled <- deck[sample(1:nrow(deck)), ]
    result1 <- play_a_round(shuffled)
    X1[i] <- result1[1]
    result2 <- play_a_round(shuffled, position=result1[2])
    X2[i] <- result2[1]
}
e1 <- mean(X1)
e1
e2 <- mean(X2)
e2
e12 <- mean(X1*X2)
```

```
e12
cov <- e12-e1*e2
cov
```

In relation to the scale of the expectations, the covariance is much smaller, with a degree of magnitude of 10^2 smaller than the other two. Hence, we can assume that the covariance is small enough to conclude X_1 and X_2 are independent. As a note, the player strategy used involved blackjack tables, in an attempt to maximize winnings.