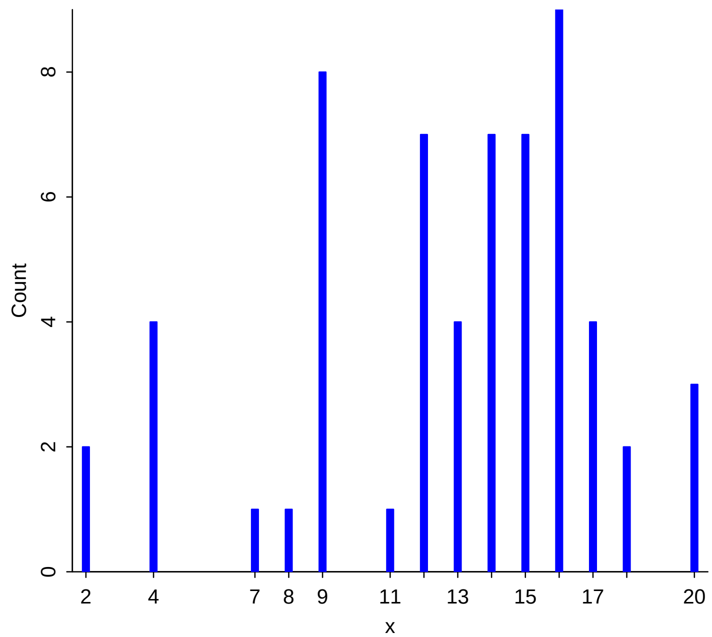
Here is a histogram of the number of times a question was answered incorrectly.



Question 16 (9 errors)

```
> muA <- 115
> sigmaA <- 15
> muB <- 100
> sigmaB<- 15
> mu.diff <- muA - muB
> sigma.diff <- sqrt(sigmaA^2 + sigmaB^2)
```

Two different normally distributed populations A and B, both have standard deviations of 15 on attribute X. However, population A has a mean of 115, while population B has a mean of 100. If you randomly sample one individual from population A and one individual from population B, what is the probability that A will have a higher X score than B? (*Hint*: Think in terms of a simple linear combination that can answer your question, and draw on the fact that linear combinations of independent normal variables are normally distributed with a variance that can be computed easily from the information given.)

Let X_a be an observation taken at random from population A, and X_b be an observation taken at random from population B. Then

$$\Pr(X_a > X_b) = \Pr(X_a - X_b > 0) \tag{1}$$

Since $D = X_a - X_b$ is a random variable, we need to know its probability distribution in order to process the problem.

From linear combination theory, and the algebra of variances and covariances covered in detail earlier in the course, we know that the mean of D is

$$\mu_a - \mu_b = 15 - 100 = 15$$

. Since the two observations are independent, the variance of D is

$$\begin{aligned} \mathrm{Var}(D) &= \mathrm{Var}(X_a - X_b) \\ &= \mathrm{Var}(X_a) + \mathrm{Var}(X_b) \\ &= 15^2 + 15^2 \\ &= 450 \end{aligned}$$

As with any linear combination of independent normal variables, D is normally distributed. So, we need the probability that D exceeds zero, which is the area to the right of zero in a normal curve with a mean of 15 and a standard deviation of $\sqrt(450)$.

```
> 1 - pnorm(0,15,sqrt(450))
[1] 0.7602499
```

Question 09 (8 errors)

Suppose that a pro basketball player is *truly* a 51.1% foul shooter, that is, the true probability of success is $p = .511$ and that foul shots are just like random binomial trials. If this player shoots 32 foul shots in one game, what is the probability that the player will be successful on 28 or more attempts?

Comment. This actually happened! Wilt Chamberlain, an NBA basketball player who had a career free-throw percentage of .511, scored 100 points in a single game. In that game, he hit 28 of 32 free throws.

With a binomial model, we could say his 32 free throws are binomial trials with $n = 32$ and $p = .511$. So what is the probability that $X \geq 28$? There are two ways (at least) to do this. One way is to use the fact that R's `dbinom` function is vectorizable, input a range of values, and sum them, as follows:

```
> sum(dbinom(28:32, 32, .511))
```

```
[1] 1.634522e-05
```

The other way is to realize that $\Pr(X \geq 28) = 1 - \Pr(X \leq 27)$, and use the `pbinom` function. This approach yields

```
> 1 - pbinom(27, 32, .511)
```

```
[1] 1.634522e-05
```

Question 12 (7 errors)

A and B are *independent* events. Suppose $\Pr(A) = 0.41$ and $\Pr B = 0.34$. Find $\Pr A \cup B$.

.

In general, $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. This is the third theorem of probability theory, which we proved in class.

However, if A and B are independent, $\Pr(A \cap B) = \Pr(A) \Pr(B)$, and so

$$\Pr(A \cup B) = 0.41 + 0.34 - (0.41)(0.34)$$

```
> 0.41 + 0.34 - 0.41*0.34
```

```
[1] 0.6106
```

Question 14 (7 errors) *Simulation and/or Analytical.* You toss a fair coin until it comes up heads for the first time. The random variable X is the number of the toss on which the first head occurred. Find the total probability that X is an even integer, i.e., $2, 4, 6, \ldots$. Which of the following is closest to the correct answer?

Let S_n be the sequence that has the first head on the nth toss. This means that there were $n-1$ tails in the preceding $n-1$ throws, and so

$$P_n = \Pr(S_n) = \Pr(X = n) = (1/2)^n$$

. The total probability of having n be even is

$$\sum_{i=1}^{\infty} P_{2i} = \sum_{i=1}^{\infty} 1/2^{2i} = \sum_{i=1}^{\infty} 1/4^i$$

This is a geometric series, and the well known result is that it converges to

$$\sum_{i=1}^{\infty} \frac{1}{4}^i = \frac{1}{4-1} = \frac{1}{3}$$

Of course, you might not know that, or might have forgotten that.
You could try simulating with R.
It is pretty clear that the answer is 1/3.

```
> p.i <- function(i){(1/4)^i}
> total.p <- function(n){sum(p.i(1:n))}
> total.p(5)

[1] 0.3330078

> total.p(10)

[1] 0.333333
```

Question 16 (7 errors) Suppose you run an opinion poll based on a random sample of size $n = 200$, asking potential voters whether they support a particular ballot initiative. If the true proportion of voters supporting the initiative is 0.50, and a binomial model is correct for the opinion poll, what is the probability that the sample proportion of voters will be within $\pm 5\%$ of the true population proportion? (*Hint.* Use the binomial distribution.)

Here, we just sum all the values (90 to 110) that produce results (0.45 to 0.55) within 5% of 50%.

```
> n <- 200
> p <- 1/2
> sum(dbinom(90:110,200,1/2))

[1] 0.8626333
```