Stat 200A, Fall 2018
A. Adhikari and Z. Bartha

WEEK 5 EXERCISES

You are expected to do all these problems, but for **Homework 5** please turn in **only Problems 2, 3, 6, and 7** on **Thursday September 27 at the start of lecture**.

1. Warmup

(a) Let H_n be the number of heads in n tosses of a coin and S_n the number of sixes in n rolls of a die. Find $E(H_n)$, $SD(H_n)$, $E(S_n)$, and $SD(S_n)$. Which is bigger: $SD(H_n)$ or $SD(S_n)$? Why?

By either the method of indicators (using the independence of the tosses / rolls) or by using part (b), $E(H_n) = \frac{n}{2}$, $SD(H_n) = \sqrt{\frac{n}{4}}$, $E(S_n) = \frac{n}{6}$, $SD(S_n) = \sqrt{\frac{5n}{36}}$.

(b) Continuing Part **a**, what are the distributions of H_n and S_n? For large n, what are the approximate distributions of H_n and S_n?

$H_n \sim Bin(n, \frac{1}{2})$, $S_n \sim Bin(n, \frac{1}{6})$. For large n these distributions are approximately $N(\frac{n}{2}, \frac{n}{4})$, and $N(\frac{n}{6}, \frac{5n}{36})$ respectively.

(c) A population consists of N elements of which G are good. Let R_n be the number of good elements in a sample of size n drawn at random with replacement from the population. Let W_n be the number of good elements in a random sample of size n drawn at random without replacement from the population. Find $E(R_n)$, $SD(R_n)$, $E(W_n)$, and $SD(W_n)$. Which is bigger: $SD(R_n)$ or $SD(W_n)$? Why?

Denoting by I_i the indicator of the event that the ith sample element is good in the first case (sampling with replacement), and by J_i in the second case (sampling without replacement), we have the following:

$$\begin{aligned}
E(R_n) &= E(I_1 + \ldots + I_n) = E(I_1) + \ldots + E(I_n) = n \cdot \frac{G}{N}, \\
Var(R_n) &= Var(I_1 + \ldots + I_n) = Var(I_1) + \ldots + Var(I_n) = n \cdot \frac{G}{N} \cdot \frac{N-G}{N}, \\
E(W_n) &= E(J_1 + \ldots + J_n) = E(J_1) + \ldots + E(J_n) = n \cdot \frac{G}{N}, \\
Var(W_n) &= Var(J_1 + \ldots + J_n) = \sum_{i=1}^{n} Var(J_i) + \sum_{i \neq j} Cov(J_i, J_j) = n \cdot \frac{G}{N} \cdot \frac{N-G}{N} + n(n-1)Cov(J_1, J_2).
\end{aligned}$$

The way to find $Cov(J_1, J_2)$ is to use the last equation for $n = N$:

$$0 = Var(W_N) = N \cdot \frac{G}{N} \cdot \frac{N-G}{N} + N(N-1)Cov(J_1, J_2),$$

which gives

$$Cov(J_1, J_2) = -\frac{G(N-G)}{N^2(N-1)}.$$

Hence,

$$Var(W_n) = n \cdot \frac{G}{N} \cdot \frac{N-G}{N} - n(n-1) \cdot \frac{G(N-G)}{N^2(N-1)} = n \cdot \frac{G}{N} \cdot \frac{N-G}{N} \cdot \left(1 - \frac{n-1}{N-1}\right).$$

This is smaller than $Var(R_n)$ by a factor of $\left(1 - \frac{n-1}{N-1}\right)$, which comes from the fact that the events that two different sample elements are good are negatively correlated.

(d) In a city that has over a million voters, 49% of the voters belong to Party A. A simple random sample of 2,500 voters is taken. Approximately what is the chance that the majority of sampled voters belong to Party A? Justify your approximation: which distribution are you approximating, and by what? Why?

The distribution of the number of voters of Party A in the sample (X) is approximately $Bin(2500, 0.49)$. This can be further approximated by the distribution $N(E(X), Var(X))$, where $E(X) = 2500 \cdot 0.49 = 1225$ and $Var(X) = 2500 \cdot 0.49 \cdot 0.51 = 624.75$, so $SD(X) \approx 25$. Hence,

$$\mathbb{P}(X > 1250) = \mathbb{P}\left(\frac{X - E(X)}{SD(X)} > \frac{1250 - E(X)}{SD(X)}\right) \approx \mathbb{P}\left(\frac{X - E(X)}{SD(X)} > \frac{1250 - 1225}{25}\right) = \mathbb{P}\left(\frac{X - E(X)}{SD(X)} > 1\right),$$

which is approximately $1 - \Phi(1)$.

2. Random Counts, Part 1

Last week you found the expectations of the random variables below. Now find the variances.

For one part you will need the fact that the SD of a geometric (p) random variable is $\frac{\sqrt{q}}{p}$ where $q = 1 - p$. We haven't proved that as the algebra takes a bit of work. We'll prove it later by conditioning.

(a) A die is rolled n times. Find the variance of number of faces that *do not* appear.

Let X be the number of faces that do not appear. With I_i being the indicator of the event that face i does not appear,

$$Var(X) = Var(I_1 + \ldots + I_6) = \sum_{i=1}^{6} Var(I_i) + \sum_{i \neq j} Cov(I_i, I_j).$$

Here $Var(I_i) = \left(\frac{5}{6}\right)^n \left(1 - \left(\frac{5}{6}\right)^n\right)$ (since $I_i \sim Bernoulli((5/6)^n)$), and

$$\begin{aligned} Cov(I_i, I_j) &= E(I_i I_j) - E(I_i)E(I_j) \\ &= \mathbb{P}(\text{neither face } i \text{ nor face } j \text{ shows up}) - \mathbb{P}(\text{face } i \text{ doesn't show up})^2 \\ &= \left(\frac{4}{6}\right)^n - \left(\frac{5}{6}\right)^{2n}. \end{aligned}$$

Hence,

$$Var(X) = 6 \cdot \left(\frac{5}{6}\right)^n \left(1 - \left(\frac{5}{6}\right)^n\right) + 6 \cdot 5 \cdot \left[\left(\frac{4}{6}\right)^n - \left(\frac{5}{6}\right)^{2n}\right].$$

(b) Use your answer to Part **a** to find the variance of the number of distinct faces that *do* appear in n rolls of a die.

Let Y be the number of faces that appear. Using the variable X from part (a), $Y = 6 - X$, thus

$$Var(Y) = Var(6 - X) = Var(-X) = Var(X).$$

(c) Find the variance of the number of times you have to roll a die till you have seen all of the faces.

Last week we showed that the waiting time can be written as $T = T_1 + T_2 + T_3 + T_4 + T_5 + T_6$, where $T_i \sim Geo\left(\frac{7-i}{6}\right)$. Moreover, these T_i's are independent. Hence,

$$Var(T) = Var(T_1 + \ldots + T_6) = Var(T_1) + \ldots + Var(T_6) = 0 + \frac{1/6}{(5/6)^2} + \frac{2/6}{(4/6)^2} + \ldots + \frac{5/6}{(1/6)^2}.$$

3. Random Counts, Part 2

(a) In the matching problem there are n letters labeled 1 through n and n envelopes labeled 1 through n. The letters are distributed at random into the envelopes, one letter per envelope, such that all $n!$ permutations are equally likely.

Let M be the number of letters that fall into envelopes with the corresponding label. That is, M is the number of "matches" or fixed points of the permutation.

Find $E(M)$ and $Var(M)$. In Week 1 Exercises, you found the approximate distribution of M for large n. Are the expectation and variance consistent with this distribution?

Let I_i be the indicator of the event that letter i falls into envelope i. I_i has the $Bernoulli(1/n)$ distribution, therefore

$$E(M) = E(I_1 + \ldots + I_n) = nE(I_1) = \frac{n}{n} = 1,$$

and

$$Var(M) = Var(I_1 + \ldots + I_n) = nVar(I_i) + n(n-1)Cov(I_1, I_2).$$

Now $Var(I_i) = \frac{1}{n} \cdot \frac{n-1}{n}$ and

$$\begin{aligned} Cov(I_1, I_2) &= E(I_1 I_2) - E(I_1)E(I_2) \\ &= \mathbb{P}(\text{both letter 1 and letter 2 fall into the corresponding envelopes}) - E(I_1)^2 \\ &= \frac{(n-2)!}{n!} - \frac{1}{n^2} = \frac{1}{n(n-1)} - \frac{1}{n^2}. \end{aligned}$$

Hence,

$$Var(M) = n \cdot \frac{1}{n} \cdot \frac{n-1}{n} + n(n-1)\left(\frac{1}{n(n-1)} - \frac{1}{n^2}\right) = 1.$$

As we saw, the approximate distribution of M for large n is $Poi(1)$, which has the exact same expectation and covariance.

(b) A deck consists of n cards, of which r are red. Cards are dealt at random without replacement till a red card appears. Let X be the number of cards dealt. Find $E(X)$ and $Var(X)$. Use symmetry; there should be no combinatorial terms or factorials in your answers.

We assume that we also deal the first red card that appears. Then if I_i is the indicator that card i (with some arbitrary numbering of the $n-r$ non-red cards) appears before the first red card, then

$$X = I_1 + I_2 + \ldots + I_{n-r} + 1.$$

The probability that a given non-red card appears before all the red cards is $\frac{1}{r+1}$ by symmetry: it appears in any of the $r+1$ possible positions relative to the red cards with equal probability. Hence, $I_i \sim Bernoulli\left(\frac{1}{r+1}\right)$. Then

$$E(X) = E(I_1 + \ldots + I_{n-r} + 1) = (n-r)E(I_1) + 1 = \frac{n-r}{r+1} + 1,$$

and

$$Var(X) = Var(I_1 + \ldots + I_{n-r} + 1) = (n-r)Var(I_1) + (n-r)(n-r-1)Cov(I_1, I_2).$$

We know that $Var(I_1) = \frac{1}{r+1} \cdot \frac{r}{r+1}$. The covariance term is the following:

$$\begin{aligned} Cov(I_1, I_2) &= E(I_1 I_2) - E(I_1)E(I_2) \\ &= \mathbb{P}(\text{both card 1 and card 2 appear before all the red cards}) - E(I_1)^2 \\ &= \frac{2r!}{(r+2)!} - \frac{1}{(r+1)^2} = \frac{2}{(r+2)(r+1)} - \frac{1}{(r+1)^2}. \end{aligned}$$

Therefore,

$$Var(X) = \frac{(n-r)r}{(r+1)^2} + (n-r)(n-r-1)\left(\frac{2}{(r+2)(r+1)} - \frac{1}{(r+1)^2}\right).$$

(c) A deck consists of n cards, of which r are red and the rest are blue. Cards are dealt at random without replacement till all the red cards have been dealt. Let Y be the number of cards dealt. Use symmetry and Part **b** to find $E(Y)$ and $Var(Y)$.

Using the variable X in part (b), Y is equal in distribution to $n - X + 1$, since they could be thought of as the number of cards in opposite parts of the deck, except for one of the red cards, which appears in both counts. Hence, $E(Y) = E(n - X + 1) = n + 1 - E(X) = n - \frac{n-r}{r+1}$ and $Var(Y) = Var(n - X + 1) = Var(X)$.

4. Correlation

The covariance of random variables X and Y has nasty units: the product of the units of X and the units of Y. Dividing the covariance by the two SDs results in an important pure number.

The *correlation coefficient* of the random variables X and Y is defined as

$$r(X,Y) \;\; = \frac{Cov(X,Y)}{SD(X)SD(Y)}$$

It is called the correlation, for short. The definition explains why X and Y are called *uncorrelated* if $Cov(X,Y) = 0$.

(a) Let X^* be X is standard units and let Y^* be Y in standard units. Check that

$$r(X,Y) = E(X^*Y^*)$$

$$E(X^*Y^*) = E\left(\frac{X - E(X)}{SD(X)} \cdot \frac{Y - E(Y)}{SD(Y)}\right) = \frac{E[(X - E(X))(Y - E(Y))]}{SD(X)SD(Y)} = \frac{Cov(X,Y)}{SD(X)SD(Y)}$$

(b) Use the fact that both $(X^* + Y^*)^2$ and $(X^* - Y^*)^2$ are non-negative random variables to show that $-1 \leq r(X,Y) \leq 1$.

[First find the numerical values of $E(X^*)$ and $E(X^{*2})$. Then find $E(X^* + Y^*)^2$.]

$E(X^*) = 0$, hence $E(X^{*2}) = Var(X^*) = 1$. Using this

$$0 \leq E(X^* + Y^*)^2 = E(X^{*2}) + E(Y^{*2}) + 2E(X^*Y^*) = 1 + 1 + 2r(X,Y),$$

hence $-1 \leq r(X,Y)$. Similarly,

$$0 \leq E(X^* - Y^*)^2 = E(X^{*2}) + E(Y^{*2}) - 2E(X^*Y^*) = 1 + 1 - 2r(X,Y),$$

hence $r(X,Y) \leq 1$.

(c) Show that if $Y = aX + b$ where $a \neq 0$, then $r(X,Y)$ is 1 or -1 depending on whether the sign of a is positive or negative.

$$r(X,Y) = \frac{Cov(X,Y)}{SD(X)SD(Y)} = \frac{Cov(X, aX+b)}{SD(X)SD(aX+b)} = \frac{aVar(X)}{SD(X)|a|SD(X)} = \frac{a}{|a|} = sign(a).$$

(d) Consider a sequence of i.i.d. Bernoulli (p) trials. For any positive integer k let X_k be the number of successes in trials 1 through k. **Use bilinearity** to find $Cov(X_n, X_{n+m})$ and hence find $r(X_n, X_{n+m})$.

Let I_i be the indicator of the event that the ith trial succeeds. Then

$$Cov(X_n, X_{n+m}) = Cov(I_1 + \ldots + I_n, I_1 + \ldots + I_{n+m}) = n \cdot Cov(I_1, I_1) = np(1-p),$$

hence

$$r(X_n, X_{n+m}) = \frac{Cov(X_n, X_{n+m})}{SD(X_n)SD(X_{n+m})} = \frac{np(1-p)}{\sqrt{np(1-p)}\sqrt{(n+m)p(1-p)}} = \sqrt{\frac{n}{n+m}}.$$

(e) Fix n and find the limit of your answer to **(d)** as $m \to \infty$. Explain why the limit is consistent with intuition.

The limit is 0, which intuitively means that if we do a large number of trials, then the outcome of the first few doesn't tell us much information about the total number of successes.

5. Relations Between Random Variables

This exercise is about departures from the "independent and identically distributed" (i.i.d.) model, with particular attention to correlation.

(a) Let X_1 and X_2 be the numbers appearing on the first and second rolls of a die. Let $S = X_1 + X_2$ and $D = X_1 - X_2$. Are S and D identically distributed? Are they independent? Are they uncorrelated?

S and D are not identically distributed, since for example S can take the value 12, while D cannot. They are not independent either, since if we know that $S = 12$ for example, then D has to be 0, while if $S = 7$, then D cannot be 0. However, they are uncorrelated, since

$$Cov(S, D) = Cov(X_1 + X_2, X_1 - X_2) = Cov(X_1, X_1) - Cov(X_2, X_2) = 0.$$

(b) Construct two random variables X and Y such that X and Y are identically distributed and negatively correlated, that is, $Cov(X, Y)$ is negative. You can do this easily on the space of a few tosses of a coin.

For example, let X be the number of heads in two tosses of a fair coin, and Y be the number of tails in the same two tosses.

(c) Construct two random variables X and Y such that $X \neq Y$, X and Y are identically distributed and positively correlated, that is, $Cov(X, Y)$ is positive. This too can be done on the space of a few tosses of a coin.

For example, let us toss a fair coin three times, and let X be the number of heads in the first two tosses and Y be the number of heads in the last two tosses.

6. The "Sample Variance"

Let $X_1, X_2, \ldots, X_n$ be i.i.d., each with mean μ and SD σ. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean.

(a) Find $E(\bar{X})$ and $SD(\bar{X})$.

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\cdot n \cdot \mu = \mu$$

$$SD(\bar{X}) = \sqrt{Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)} = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)} = \sqrt{\frac{1}{n^2}\cdot n\sigma^2} = \frac{\sigma}{\sqrt{n}}$$

(b) For each i, find $Cov(X_i, \bar{X})$. [Plug in the definition of $\bar{X}$ and use bilinearity.]

$$Cov(X_i, \bar{X}) = Cov\left(X_i, \frac{1}{n}\sum_{i=1}^{n} X_i\right) = Cov\left(X_i, \frac{1}{n}X_i\right) = \frac{\sigma^2}{n}$$

(c) For each i in the range 1 through n, define the *ith deviation in the sample* as $D_i = X_i - \bar{X}$. Find $E(D_i)$ and $Var(D_i)$. [Write the variance as $Cov(D_i, D_i)$, plug in the definition of D_i, and use bilinearity.]

$$E(D_i) = E(X_i - \bar{X}) = E(X_i) - E(\bar{X}) = \mu - \mu = 0,$$

and

$$\begin{aligned} Var(D_i) &= Cov(D_i, D_i) = Cov(X_i - \bar{X}, X_i - \bar{X}) = Var(X_i) + Var(\bar{X}) - 2Cov(X_i, \bar{X}) \\ &= \sigma^2 + \frac{\sigma^2}{n} - 2\cdot\frac{\sigma^2}{n} = \sigma^2\left(1 - \frac{1}{n}\right). \end{aligned}$$

(d) Define the random variable $\hat{\sigma}^2$ as

$$\hat{\sigma}^2 \;=\; \frac{1}{n}\sum_{i=1}^{n} D_i^2$$

Find $E(\hat{\sigma}^2)$.

For this random variable, the notation $\hat{\sigma}^2$ is pretty standard in statistics. Just think of $\hat{\sigma}^2$ as a symbol; it doesn't help to start thinking about the random variable that is its square root.

$$E(\hat{\sigma}^2) = \frac{1}{n}\sum_{i=1}^{n} E(D_i^2) = \frac{1}{n}\sum_{i=1}^{n}\sigma^2\left(1 - \frac{1}{n}\right) = \sigma^2\left(1 - \frac{1}{n}\right)$$

(e) Use Part **d** to construct a random variable denoted S^2 that is an unbiased estimator of σ^2. This random variable S^2 is called the *sample variance*.

Since $E(\hat{\sigma}^2) = \sigma^2\left(1 - \frac{1}{n}\right) = \sigma^2 \cdot \frac{n-1}{n}$, multiplying both sides by $\frac{n}{n-1}$ gives that $S^2 = \frac{n}{n-1}\cdot\hat{\sigma}^2$ is an unbiased estimator of σ^2.

7. Geometric Mean

(a) Let U have the uniform distribution on the interval $(0, 1)$. Find the cdf of $-\log(U)$. Identify this as the cdf of a well known distribution and provide the relevant parameters.

The range of $-\log(U)$ is $(0,\infty)$. For $z>0$,

$$\mathbb{P}(-\log(U)<z)=\mathbb{P}(\log(U)>-z)=\mathbb{P}(U>e^{-z})=1-e^{-z}.$$

Hence, $-\log(U)\sim Exp(1)$.

(b) Let $U_1,U_2,\ldots,U_n$ be i.i.d. uniform on the interval $(0,1)$, and let $G_n=(U_1U_2\cdots U_n)^{1/n}$ be the geometric mean of the sample. Show that there is a constant c such that $G_n\xrightarrow{P}c$ as $n\to\infty$, and find c.

$\log G_n=\frac{1}{n}\sum_{i=1}^n\log(U_i)$, so

$$E(\log G_n)=\frac{1}{n}\sum_{i=1}^n E(\log(U_i))=E(\log(U_1))=-1,$$
$$Var(\log G_n)=\frac{1}{n^2}\sum_{i=1}^n Var(\log(U_i))=\frac{1}{n}Var(\log(U_1))=\frac{1}{n}.$$

Since $Var(\log G_n)\to 0$ as $n\to\infty$ and $E(\log G_n)$ is constant -1, we have that $\log G_n\xrightarrow{P}-1$, so $G_n\xrightarrow{P}e^{-1}$.

(c) For large n and small $\epsilon>0$, approximate $P(|G_n-c|<\epsilon)$. Justify your answer.

$$\begin{aligned}
\mathbb{P}(|G_n-c|<\epsilon)&=\mathbb{P}(-\epsilon<G_n-e^{-1}<\epsilon)=\mathbb{P}(e^{-1}-\epsilon<G_n<e^{-1}+\epsilon)\\
&=\mathbb{P}(\log(e^{-1}-\epsilon)<\log G_n<\log(e^{-1}+\epsilon))\\
&\approx\mathbb{P}\left(\log(e^{-1})-\frac{\epsilon}{e^{-1}}<\log G_n<\log(e^{-1})+\frac{\epsilon}{e^{-1}}\right)\\
&=\mathbb{P}(-1-\epsilon e<\log G_n<-1+\epsilon e)\\
&=\mathbb{P}\left(-\epsilon e\sqrt{n}<\frac{\log G_n+1}{1/\sqrt{n}}<\epsilon e\sqrt{n}\right)\\
&\approx\Phi(\epsilon e\sqrt{n})-\Phi(-\epsilon e\sqrt{n})=1-2\Phi(-\epsilon e\sqrt{n}),
\end{aligned}$$

where we used that $\frac{\log G_n-E(\log G_n)}{SD(\log G_n)}=\frac{\log G_n-(-1)}{1/\sqrt{n}}$ is approximately standard normal, since $\log G_n=\frac{1}{n}\sum_{i=1}^n\log(U_i)$, the average of i.i.d. variables.

8. Empty Boxes

There are n balls and $2n$ boxes. Each ball is placed in a box picked uniformly at random, independent of the placement of all other balls. Let W_n be the proportion of empty boxes.

(a) Find $E(W_n)$ and $Var(W_n)$.

Let I_i be the indicator that the ith box is empty. Then $I_i\sim Bernoulli\left(\left(\frac{2n-1}{2n}\right)^n\right)$ and the number of empty boxes is $2nW_n=I_1+\ldots+I_{2n}$. Hence,

$$E(W_n)=\frac{1}{2n}E(2nW_n)=\frac{1}{2n}E(I_1+\ldots+I_{2n})=\frac{2n}{2n}E(I_1)=\left(\frac{2n-1}{2n}\right)^n,$$

and

$$\begin{aligned}
Var(W_n)&=\frac{1}{(2n)^2}Var(2nW_n)=\frac{1}{(2n)^2}Var(I_1+\ldots+I_{2n})\\
&=\frac{1}{(2n)^2}\big[2nVar(I_1)+2n(2n-1)Cov(I_1,I_2)\big].
\end{aligned}$$

Here $Var(I_1) = \left(\frac{2n-1}{2n}\right)^n \left(1 - \left(\frac{2n-1}{2n}\right)^n\right)$, and

$$\begin{aligned} Cov(I_1, I_2) &= E(I_1 I_2) - E(I_1)E(I_2) \\ &= \mathbb{P}(\text{both box 1 and box 2 are empty}) - E(I_1)^2 \\ &= \left(\frac{2n-2}{2n}\right)^n - \left(\frac{2n-1}{2n}\right)^{2n}. \end{aligned}$$

Therefore,

$$Var(W_n) = \frac{1}{(2n)^2}\left[2n\left(\frac{2n-1}{2n}\right)^n\left(1 - \left(\frac{2n-1}{2n}\right)^n\right) + 2n(2n-1)\left(\left(\frac{2n-2}{2n}\right)^n - \left(\frac{2n-1}{2n}\right)^{2n}\right)\right].$$

(b) Show that there is a constant c such that $W_n \xrightarrow{P} c$ as $n \to \infty$, and find c.

Using the fact that $(1 + x/n)^n \to e^x$ for any x, we can find the limit of both $E(W_n)$ and $Var(W_n)$:

$$\lim_{n\to\infty} E(W_n) = \lim_{n\to\infty}\left(1 - \frac{1}{2n}\right)^n = e^{-1/2},$$

and

$$\lim_{n\to\infty} Var(W_n) = 0$$

after carefully looking at each term. Thus, $W_n \xrightarrow{P} e^{-1/2}$.

9. Reliability

Let X_n be the number of successes in n i.i.d. Bernoulli (0.9) trials. About how large does n have to be so that the chance of 100 or more successes is about 99%?

Versions of this calculation are used by airlines to work out by how much they will overbook their flights, or by manufacturers who need to get a minimum number of good items using a process that has some chance of producing duds.

Using that $X_n \sim Binomial(n, 0.9)$ and its normal approximation:

$$\mathbb{P}(X_n \geq 100) = \mathbb{P}\left(\frac{X_n - E(X_n)}{SD(X_n)} \geq \frac{100 - E(X_n)}{SD(X_n)}\right) = \mathbb{P}\left(\frac{X_n - 0.9n}{\sqrt{0.09n}} \geq \frac{100 - 0.9n}{\sqrt{0.09n}}\right) \approx 1 - \Phi\left(\frac{100 - 0.9n}{\sqrt{0.09n}}\right).$$

We want this to be 0.99, so $\Phi\left(\frac{100-0.9n}{\sqrt{0.09n}}\right) = 0.01$, which gives $\frac{100-0.9n}{\sqrt{0.09n}} \approx -2.3$. Solving this for n, we get $n \approx 120$.