Section 8 – Solutions

Review

1) Multivariate: Discrete to Continuous:

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = \mathbb{P}(X=x, Y=y)$	$f_{X,Y}(x,y) \neq \mathbb{P}(X=x, Y=y)$
Joint range/support $\Omega_{X,Y}$	$\{(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) > 0\}$	$\{(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) > 0\}$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(t,s)\,ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
Expectation	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y)p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dxdy$
Independence must have	$\forall x,y, p_{X,Y}(x,y) = p_X(x)p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x,y, f_{X,Y}(x,y) = f_X(x)f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$
Conditional PMF/PDF	$p_{X\|Y}(x\|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X\|Y=y] = \sum_x x \cdot p_{X\|Y}(x\|y)$	$\mathbb{E}[X\|Y=y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y)dx$

2) Normal (Gaussian, "bell curve"): $X \sim \mathcal{N}(\mu, \sigma^2)$ iff X has the following probability density function:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \quad x \in \mathbb{R}$$

$\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$. The "standard normal" random variable is typically denoted Z and has mean 0 and variance 1: if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. The CDF has no closed form, but we denote the CDF of the standard normal as $\Phi(z) = F_Z(z) = \mathbb{P}(Z \leq z)$. Note from symmetry of the probability density function about $z = 0$ that: $\Phi(-z) = 1 - \Phi(z)$.

3) Central Limit Theorem (CLT):

Let $X_1, \ldots, X_n$ be iid random variables with $\mathbb{E}[X_i] = \mu$ and $Var(X_i) = \sigma^2$. Let $X = \sum_{i=1}^n X_i$, which has $\mathbb{E}[X] = n\mu$ and $Var(X) = n\sigma^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, which has $\mathbb{E}[\overline{X}] = \mu$ and $Var(\overline{X}) = \frac{\sigma^2}{n}$. $\overline{X}$ is called the *sample mean*. Then, as $n \to \infty$, $\overline{X}$ approaches the normal distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Standardizing, this is equivalent to $Y = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ approaching $\mathcal{N}(0,1)$. Similarly, as $n \to \infty$, X approaches $\mathcal{N}(n\mu, n\sigma^2)$ and $Y' = \frac{X - n\mu}{\sigma\sqrt{n}}$ approaches $\mathcal{N}(0,1)$.

It is no surprise that $\overline{X}$ has mean μ and variance σ^2/n – this can be done with simple calculations. The importance of the CLT is that, for large n, regardless of what distribution X_i comes from, $\overline{X}$ is *approximately normally distributed with mean* μ *and variance* σ^2/n. Don't forget the continuity correction, only when $X_1, \ldots, X_n$ are discrete random variables.

Here is the **Standard normal table.**

4) Uniform: $X \sim \text{Uniform}(a,b)$ iff X has the following probability density function:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{a+b}{2}$ and $\mathrm{Var}\left(()\,X\right) = \frac{(b-a)^2}{12}$. This represents each real number from $[a,b]$ to be equally likely.

5) Exponential: $X \sim \text{Exponential}(\lambda)$ iff X has the following probability density function:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$. $F_X(x) = 1 - e^{-\lambda x}$ for $x \geqslant 0$. The exponential random variable is the continuous analog of the geometric random variable: it represents the waiting time to the next event, where $\lambda > 0$ is the average number of events per unit time. Note that the exponential measures how much time passes until the next event (any real number, continuous), whereas the Poisson measures how many events occur in a unit of time (nonnegative integer, discrete). The exponential random variable X is memoryless:

$$\text{for any } s, t \geqslant 0,\ \mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$$

The geometric random variable also has this property.

Task 1 – Joint PMF's

Suppose X and Y have the following joint PMF:

X/Y	1	2	3
0	0	0.2	0.1
1	0.3	0	0.4

a) Identify the range of X (Ω_X), the range of Y (Ω_Y), and their joint range ($\Omega_{X,Y}$).

$\Omega_X = \{0, 1\}$, $\Omega_Y = \{1, 2, 3\}$, and $\Omega_{X,Y} = \{(0,2), (0,3), (1,1), (1,3)\}$

b) Find the marginal PMF for X, $p_X(x)$ for $x \in \Omega_X$.

$$p_X(0) = \sum_y p_{X,Y}(0, y) = 0 + 0.2 + 0.1 = 0.3$$

$$p_X(1) = 1 - p_X(0) = 0.7$$

c) Find the marginal PMF for Y, $p_Y(y)$ for $y \in \Omega_Y$.

$$p_Y(1) = \sum_x p_{X,Y}(x, 1) = 0 + 0.3 = 0.3$$

$$p_Y(2) = \sum_x p_{X,Y}(x, 2) = 0.2 + 0 = 0.2$$

$$p_Y(3) = \sum_x p_{X,Y}(x, 3) = 0.1 + 0.4 = 0.5$$

d) Are X and Y independent? Why or why not?

No, since a necessary condition is that $\Omega_{X,Y} = \Omega_X \times \Omega_Y$.

e) Find $\mathbb{E}[X^3 Y]$.

Note that $X^3 = X$ since X takes values in $\{0, 1\}$.

$$\mathbb{E}[X^3 Y] = \mathbb{E}[XY] = \sum_{(x,y) \in \Omega_{X,Y}} xy p_{X,Y}(x, y) = 1 \cdot 1 \cdot 0.3 + 1 \cdot 3 \cdot 0.4 = 1.5$$

Task 2 – Do You "Urn" to Learn More About Probability?

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let $X_i = 1$ if the i-th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

a) X_1, X_2

Here is one way of defining the joint pmf of X_1, X_2

$$p_{X_1,X_2}(1,1) = \mathbb{P}(X_1 = 1)\,\mathbb{P}(X_2 = 1 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{4}{12} = \frac{20}{156}$$

$$p_{X_1,X_2}(1,0) = \mathbb{P}(X_1 = 1)\,\mathbb{P}(X_2 = 0 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{8}{12} = \frac{40}{156}$$

$$p_{X_1,X_2}(0,1) = \mathbb{P}(X_1 = 0)\,\mathbb{P}(X_2 = 1 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{5}{12} = \frac{40}{156}$$

$$p_{X_1,X_2}(0,0) = \mathbb{P}(X_1 = 0)\,\mathbb{P}(X_2 = 0 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{7}{12} = \frac{56}{156}$$

b) X_1, X_2, X_3

Instead of listing out all the individual probabilities, we could write a more compact formula for the pmf. In this problem, the denominator is always $P(13, k)$, where k is the number of random variables in the joint pmf. And the numerator is $P(5, i)$ times $P(8, j)$ where i and j are the number of 1s and 0s, respectively.

If we wish to compute $p_{X_1,X_2,X_3}(x_1, x_2, x_3)$, then the number of 1s (i.e., white balls) is $x_1+x_2+x_3$, and the number of 0s (i.e., red balls) is $(1 - x_1) + (1 - x_2) + (1 - x_3)$. Then, we can write the pmf as follows:

$$p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \frac{10!}{13!} \cdot \frac{5!}{(5 - x_1 - x_2 - x_3)!} \cdot \frac{8!}{(5 + x_1 + x_2 + x_3)!}$$

Task 3 – Trinomial Distribution

A generalization of the Binomial model is when there is a sequence of n independent trials, but with three outcomes, where $\mathbb{P}(\text{outcome } i) = p_i$ for $i = 1, 2, 3$ and of course $p_1 + p_2 + p_3 = 1$. Let X_i be the number of times outcome i occurred for $i = 1, 2, 3$, where $X_1 + X_2 + X_3 = n$. Find the joint PMF $p_{X_1,X_2,X_3}(x_1, x_2, x_3)$ and specify its value for all $x_1, x_2, x_3 \in \mathbb{R}$.

Are X_1 and X_2 independent?

Same argument as for the binomial PMF:

$$p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3} \prod_{i=1}^{3} p_i^{x_i} = \frac{n!}{x_1! x_2! x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

where $x_1 + x_2 + x_3 = n$ and are nonnegative integers.

X_1 and X_2 are not independent. For example $Pr(X_1 = n) > 0$ and $Pr(X_2 = n) > 0$, but $Pr(X_1 = n, X_2 = n) = 0$.

Task 4 – Successes

Consider a sequence of independent Bernoulli trials, each of which is a success with probability p. Let X_1 be the number of failures preceding the first success, and let X_2 be the number of failures between the first 2 successes. Find the joint pmf of X_1 and X_2. Write an expression for $E[\sqrt{X_1 X_2}]$. You can leave your answer in the form of a sum.

X_1 and X_2 take on two particular values x_1 and x_2, when there are x_1 failures followed by one success, and then x_2 failures followed by one success. Since the Bernoulli trials are independent the joint pmf is

$$p_{X_1,X_2}(x_1,x_2) = (1-p)^{x_1} p \cdot (1-p)^{x_2} p = (1-p)^{x_1+x_2} p^2$$

for $(x_1, x_2) \in \Omega_{X_1,X_2} = \{0,1,2,\ldots\} \times \{0,1,2,\ldots\}$. By the definition of expectation

$$E[\sqrt{X_1 X_2}] = \sum_{(x_1,x_2)\in\Omega_{X_1,X_2}} \sqrt{x_1 x_2} \cdot (1-p)^{x_1+x_2} p^2.$$

Task 5 – Who fails first?

Here's a question that commonly comes up in industry, but isn't immediately obvious. You have a disk with probability p_1 of failing each day. You have a CPU which independently has probability p_2 of failing each day. What is the probability that your disk fails *before* your CPU?

a) Compute the probability by summing over the relevant part of the probability space.

We model the problem by considering two Geometric random variables and deriving the probability that one is smaller than the other. Let $X_1 \sim$ Geometric(p_1). Let $X_2 \sim$ Geometric(p_2). Assume X_1 and X_2 are independent. We want $\mathbb{P}(X_1 < X_2)$.

$$\begin{aligned}
\mathbb{P}(X_1 < X_2) &= \sum_{k=1}^{\infty} \sum_{k_2=k+1}^{\infty} p_{X_1,X_2}(k, k_2) \\
&= \sum_{k=1}^{\infty} \sum_{k_2=k+1}^{\infty} p_{X_1}(k) \cdot p_{X_2}(k_2) && \text{(by independence)} \\
&= \sum_{k=1}^{\infty} \sum_{k_2=k+1}^{\infty} (1-p_1)^{k-1} p_1 \cdot (1-p_2)^{k_2-1} p_2 \\
&= \sum_{k=1}^{\infty} (1-p_1)^{k-1} p_1 \sum_{k_2=k+1}^{\infty} (1-p_2)^{k_2-1} p_2 \\
&= \sum_{k=1}^{\infty} (1-p_1)^{k-1} p_1 (1-p_2)^{k} \sum_{k_2=1}^{\infty} (1-p_2)^{k_2-1} p_2 \\
&= \sum_{k=1}^{\infty} (1-p_1)^{k-1} p_1 (1-p_2)^{k} \cdot 1 \\
&= p_1(1-p_2) \sum_{k=1}^{\infty} [(1-p_2)(1-p_1)]^{k-1} \\
&= \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}.
\end{aligned}$$

b) Try to provide an intuitive reason for the answer.

Think about X_1 and X_2 in terms of coin flips. Notice that all the flips are irrelevant until the final flip, since before the final flip, both the X_1 coin and the X_2 coin only yield tails. $\mathbb{P}(X_1 < X_2)$ is the probability that on the final flip, where by definition at least one coin comes up heads, it is the case that the X_1 coin is heads and the X_2 coin is tails. So we're looking for the probability that the X_1 coin produces a heads and the X_2 coin produces a tails, conditioned on the fact that they're not both tails, which is derived as:

$$\begin{aligned}\mathbb{P}(\text{Coin 1 } = H \& \text{ Coin 2 } = T \mid \text{not both } T) &= \frac{\mathbb{P}(\text{Coin 1 } = H \& \text{ Coin 2 } = T)}{\mathbb{P}(\text{not both } T)} \\ &= \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}.\end{aligned}$$

Another way to approach this problem is to use conditioning. Recall that in computing the probability of an event, we saw in Chapter 2 that it is often useful to condition on other events. We can use this same idea in computing probabilities involving random variables, because $X = k$ and $Y = y$ are just events.

c) Recompute the probability using the law of total probability, conditioning on the value of X_1.

Again, let $X_1 \sim \text{Geometric}(p_1)$ and $X_2 \sim \text{Geometric}(p_2)$, where X_1 and X_2 are independent. Then

$$\begin{aligned}\mathbb{P}(X_1 < X_2) &= \sum_{k=1}^{\infty} \mathbb{P}(X_1 < X_2 \mid X_1 = k) \cdot \mathbb{P}(X_1 = k) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(k < X_2 \mid X_1 = k) \cdot \mathbb{P}(X_1 = k) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(X_2 > k) \cdot \mathbb{P}(X_1 = k) && \text{(by independence)} \\ &= \sum_{k=1}^{\infty} (1-p_2)^k \cdot (1-p_1)^{k-1} \cdot p_1 \\ &= p_1(1-p_2) \sum_{k=1}^{\infty} [(1-p_2)(1-p_1)]^{k-1} \\ &= \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}.\end{aligned}$$

Task 6 – Continuous joint density

The joint density of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of W and V is given by

$$f_{W,V}(w,v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

For two random variables X, Y to be independent, we must have $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ for all $x \in \Omega_X,\ y \in \Omega_Y$. Let's start with X and Y by finding their marginal PDFs. By definition, and using the fact that the joint PDF is 0 outside of $y > 0$, we get:

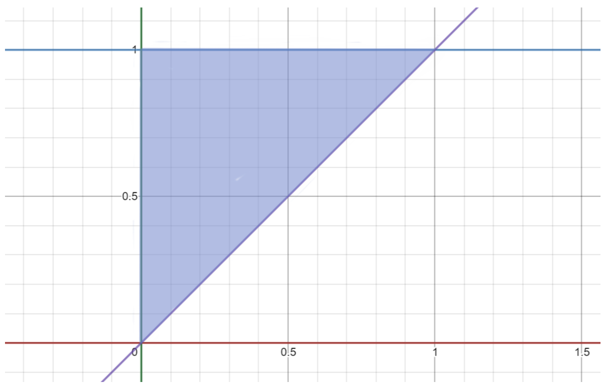
$$f_X(x) = \int_0^\infty xe^{-(x+y)} dy = e^{-x}x$$

We do the same to get the PDF of Y, again over the range $x > 0$:

$$f_Y(y) = \int_0^\infty xe^{-(x+y)} dx = e^{-y}$$

Since $e^{-x}x \cdot e^{-y} = xe^{-x-y} = xe^{-(x+y)}$ for all $x, y > 0$, X and Y are independent.

We can see that W and V are not independent simply by observing that $\Omega_W = (0, 1)$ and $\Omega_V = (0, 1)$, but $\Omega_{W,V}$ is not equal to their Cartesian product. Specifically, looking at their range of $f_{W,V}(w, v)$. Graphing it with w as the "x-axis" and v as the "y-axis", we see that :



The shaded area is where the joint pdf is strictly positive. Looking at it, we can see that it is not rectangular, and therefore it is not the case that $\Omega_{W,V} = \Omega_W \times \Omega_V$. Remember, the joint range being the Cartesian product of the marginal ranges is not sufficient for independence, but it is *necessary*. Therefore, this is enough to show that they are not independent.

Task 7 – Law of Total Probability Review

a) (Discrete version) Suppose we flip a coin with probability U of heads, where U is equally likely to be one of $\Omega_U = \{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$ (notice this set has size $n+1$). Let H be the event that the coin comes up heads. What is $\mathbb{P}(H)$?

We can use the law of total probability, conditioning on $U = \frac{k}{n}$ for $k = 0, ..., n$. Note that the probability of getting heads conditioning on a fixed U value is U, and that the probability of U

taking on any value in its range is $\frac{1}{n+1}$ since it is discretely uniform.

$$\begin{aligned}\mathbb{P}(H) &= \sum_{k=0}^{n} \mathbb{P}\left(H \mid U = \frac{k}{n}\right)\mathbb{P}\left(U = \frac{k}{n}\right) \\ &= \sum_{k=0}^{n} \frac{k}{n}\cdot\frac{1}{n+1} \\ &= \frac{1}{n(n+1)}\sum_{k=0}^{n} k \\ &= \frac{1}{n(n+1)}\frac{n(n+1)}{2} = \frac{1}{2}\end{aligned}$$

b) Now suppose $U \sim$ Uniform(0,1) has the *continuous* uniform distribution over the interval $[0, 1]$. What is $\mathbb{P}(H)$?

Use the continuous version of the law of total probability: suppose E is an event, and X is a continuous random variable with density function $f_X(x)$. Then

$$\mathbb{P}(E) = \int_{-\infty}^{\infty} \mathbb{P}(E \mid X = x) f_X(x)dx$$

We do the same thing, this time using the continuous law of total probability. Note, this time, that we're conditioning on $U = u$ and taking the integral with respect to u, and that the density of U for any value in its range is 1 because it is uniformly random.

$$\mathbb{P}(H) = \int_{-\infty}^{\infty} \mathbb{P}(H \mid U = u) f_U(u)du$$

We can take the integral from 0 to 1 instead because outside of that range the density of U is 0.

$$= \int_0^1 \mathbb{P}(H \mid U = u) f_U(u)du = \int_0^1 u \cdot 1\, du = \frac{1}{2}[u^2]_0^1 = \frac{1}{2}$$

Task 8 – Normal Approximation of a Sum

Imagine that we are trying to transmit a signal. During the transmission, there are a hundred sources independently making low noise. Each source produces an amount of noise that is Uniformly distributed between $a = -1$ and $b = 1$. If the total amount of noise is greater than 10 or less than -10, then it corrupts the signal. However, if the absolute value of the total amount of noise is under 10, then it is not a problem. What is the approximate probability that the absolute value of the total amount of noise from the 100 signals is less than 10?

Let S be the total amount of noise. We want to find $\mathbb{P}(|S| < 10) = \mathbb{P}(-10 < S < 10)$. Let X_i be the noise from source i. Then, we have

$$S = \sum_{i=1}^{100} X_i.$$

Since the X_i are uniformly distributed, we have that $\mathbb{E}[X_i] = \frac{a+b}{2} = 0$ and $\text{Var}(X_i) = \frac{(b-a)^2}{12} = \frac{1}{3}$. Since the X_i are i.i.d, by the Central Limit Theorem, we find that S is approximately distributed according to $N\left(0, 100 \cdot \frac{1}{3}\right)$. Now, we standardize to get

$$\begin{aligned}\mathbb{P}(-10 < S < 10) &= \mathbb{P}\left(\frac{-10-0}{\sqrt{100/3}} < \frac{S-0}{\sqrt{100/3}} < \frac{10-0}{\sqrt{100/3}}\right) \\ &= 2\Phi(\sqrt{3}) - 1 \approx 0.91\end{aligned}$$

Task 9 – Confidence Intervals

Suppose that $X_1, \ldots, X_n$ are i.i.d. samples from a normal distribution with unknown mean μ and variance 36. How big does n need to be so that μ is in

$$[\overline{X} - 0.11, \overline{X} + 0.11]$$

with probability at least 0.97?

Recall that

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

You may use the fact that $\Phi^{-1}(0.985) = 2.17$.

Our goal is to find n such that μ lies within 0.11 of $\bar{X}$ 97% of the time. This is equivalent to finding n such that the probability that μ lies outside the range is less than 3%.

$$\mathbb{P}(|\bar{X} - \mu| > 0.11) \leqslant 0.03$$

Let us define $Z = \frac{\bar{X}-\mu}{\sigma}$. We can solve for σ by using the Properties of Variance. Since

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

we can say that

$$\text{Var}(\bar{X}) = \text{Var}(\frac{1}{n}\sum_{i=1}^{n} X_i)$$

Using the Properties of Variance and the fact that X_i's are i.i.d., $\text{Var}(\bar{X}) = \frac{1}{n^2} \cdot n \cdot 36 = \frac{36}{n}$, so $\sigma = \frac{6}{\sqrt{n}}$.

$$\begin{aligned}
\mathbb{P}(|\bar{X} - \mu| > 0.11) &\leqslant 0.03 \\
\mathbb{P}(|Z| \cdot \sigma > 0.11) &\leqslant 0.03 && \text{[Definition of } Z\text{]} \\
\mathbb{P}\left(|Z| > \frac{0.11}{6}\sqrt{n}\right) &\leqslant 0.03 \\
\mathbb{P}\left(Z < -\frac{0.11}{6}\sqrt{n}\right) &\leqslant 0.015 && \text{[Symmetry of Normal Dist.]} \\
\Phi\left(-\frac{0.11}{6}\sqrt{n}\right) &\leqslant 0.015 && \text{[CDF of Standard Norm.]} \\
-\frac{0.11}{6}\sqrt{n} &\leqslant -\Phi^{-1}(0.985) \\
\sqrt{n} &\geqslant \frac{6 \cdot \Phi^{-1}(0.985)}{0.11} \\
n &\geqslant \left(\frac{6 \cdot \Phi^{-1}(0.985)}{0.11}\right)^2 \\
&\approx 14009.95
\end{aligned}$$

Then n must be at least 14010.