MAST20004 Probability

Assignment Four: Solutions

Problem 1. Let

$$P(z) = c(z+1)(z+2)(2z+1), \quad z \in \mathbb{R}$$

be the probability generating function of a random variable X.

(i) Find the constant c.
(ii) Find the pmf of X.
(iii) Compute the mean and variance of X by using the function $P(z)$.
(iv) Show that X has the same distribution as the sum of three independent Bernoulli random variables.

Solution. (i) We expand $P(z)$ as

$$P(z) = c \times (2z^3 + 7z^2 + 7z + 2).$$

Since $P(z)$ is a pgf, the coefficients must add up to one:

$$c \times (2+7+7+2) = 1 \implies c = \frac{1}{18}.$$

One can also use $P(1) = 1$ to obtain $c = 1/18$.
(ii) The pmf of X is given by

$$p(0) = \frac{2}{18},\ p(1) = \frac{7}{18},\ p(2) = \frac{7}{18},\ p(3) = \frac{2}{18}.$$

(iii) We first compute the derivatives of $P(z)$:

$$P'(z) = \frac{6z^2 + 14z + 7}{18},\ P''(z) = \frac{12z + 14}{18}.$$

Therefore,

$$\mathbb{E}[X] = P'(1) = \frac{3}{2},$$

and

$$V[X] = P''(1) + P'(1) - P'(1)^2 = \frac{25}{36}.$$

(iv) We can rearrange $P(z)$ as

$$P(z) = \frac{(z+1)(z+2)(2z+1)}{18} = \big(\frac{1}{2} + \frac{z}{2}\big) \times \big(\frac{2}{3} + \frac{z}{3}\big) \times \big(\frac{1}{3} + \frac{2z}{3}\big).$$

This is the product of the pgf's of three Bernoulli random variables with parameters $1/2, 1/3, 2/3$ respectively. According to the convolution theorem and the uniqueness theorem, we conclude that X has the same distribution as the sum of independent $B(1, 1/2)$-, $B(1, 1/3)$-, and $B(1, 2/3)$-random variables.

Problem 2. Let X be a continuous uniform random variable over $(0, \pi)$.

(i) Let $Y = \sin X$. Use the second order Taylor expansion of $\psi(x) = \sin x$ to approximate $\mathbb{E}[Y]$.
(ii) Let $X_1, \cdots, X_{100}$ be independent random variables, all having the same distribution as X does. Let $S_{100} = X_1 + \cdots + X_{100}$. Use Chebyshev's inequality to estimate the probability

$$\mathbb{P}(|S_{100} - \mathbb{E}[S_{100}]| \geqslant 20).$$

(iii) Use the central limit theorem to approximate the same probability as in Part (ii).

Solution. (i) We have

$$\mu = \frac{\pi}{2}, \ \sigma^2 = \frac{\pi^2}{12},$$

and

$$\psi(x) = \sin x, \ \psi''(x) = -\sin x.$$

Therefore,

$$\begin{aligned}\mathbb{E}[Y] &\approx \psi(\mu) + \frac{1}{2}\psi''(\mu)\sigma^2 = \sin\frac{\pi}{2} - \frac{1}{2}\big(\sin\frac{\pi}{2}\big) \times \frac{\pi^2}{12} \\ &= 1 - \frac{\pi^2}{24} \approx 0.58877.\end{aligned}$$

(ii) According to Chebyshev's inequality, we have

$$\mathbb{P}(|S_{100} - \mathbb{E}[S_{100}]| \geqslant 20) \leqslant \frac{1}{400}V[S_{100}] = \frac{1}{400} \times 100 \times \frac{\pi^2}{12} \approx 0.20562.$$

(iii) According to the central limit theorem, we have

$$\begin{aligned}\mathbb{P}(|S_{100}-\mathbb{E}[S_{100}]|\geqslant 20) &= \mathbb{P}\Big(\Big|\frac{S_{100}-\mathbb{E}[S_{100}]}{\sqrt{V[S_{100}]}}\geqslant \frac{20}{\sqrt{100\times\frac{\pi^2}{12}}}\Big|\Big)\\ &\approx \mathbb{P}(|Z|\geqslant 2.21)\quad (Z\sim N(0,1))\\ &= 2\times(1-\Phi(2.21))\\ &\approx 0.0271.\end{aligned}$$

Problem 3. Let $\{X_n : n\geqslant 1\}$ be an independent sequence of Bernoulli random variables with parameter $p=1/2$. Let $S_n = X_1+\cdots+X_n$.

(i) By directly computing moment generating functions, show that $S_n \stackrel{d}{=} B(n,1/2)$.
(ii) By using the method of moment generating functions, show that

$$\frac{S_n-\mathbb{E}[S_n]}{\sqrt{V[S_n]}} \xrightarrow{d} N(0,1)$$

as $n\to\infty$.

Solution. (i) The mgf of each X_i is given by

$$M_{X_i}(t)=\mathbb{E}[e^{tX_i}]=e^{t\times 0}\times\frac{1}{2}+e^{t\times 1}\times\frac{1}{2}=\frac{1+e^t}{2}.$$

Since $X_1,\cdots,X_n$ are independent, we have

$$M_{S_n}(t)=\mathbb{E}\big[e^{t(X_1+\cdots+X_n)}\big]=\mathbb{E}[e^{tX_1}]\cdots\mathbb{E}[e^{tX_n}]=\frac{(1+e^t)^n}{2^n}.$$

On the other hand, the mgf of $Y\stackrel{d}{=}B(n,1/2)$ is given by

$$\begin{aligned}M_Y(t)=\mathbb{E}[e^{tY}] &= \sum_{k=0}^n e^{tk}\cdot\mathbb{P}(Y=k)\\ &= \sum_{k=0}^n e^{tk}\cdot\binom{n}{k}\big(\frac{1}{2}\big)^n=\frac{1}{2^n}\sum_{k=0}^n\binom{n}{k}(e^t)^k\cdot 1^{n-k}\\ &= \frac{(1+e^t)^n}{2^n}\quad\text{(the binomial theorem)}.\end{aligned}$$

By the uniqueness theorem for the mgf, we conclude that $S_n\stackrel{d}{=}B(n,1/2)$.
(ii) First of all, we have

$$\mathbb{E}[S_n]=\frac{n}{2},\ V[S_n]=\frac{n}{4}.$$

The mgf of

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{V[S_n]}} = \frac{2S_n - n}{\sqrt{n}}$$

is given by

$$\begin{aligned} M_n(t) = \mathbb{E}[e^{t\left(\frac{2S_n-n}{\sqrt{n}}\right)}] &= e^{-t\sqrt{n}} \cdot M_{S_n}\big(\frac{2t}{\sqrt{n}}\big) \\ &= e^{-t\sqrt{n}} \cdot \frac{(1+e^{2t/\sqrt{n}})^n}{2^n} \\ &= \big(\frac{e^{-t/\sqrt{n}} + e^{t/\sqrt{n}}}{2}\big)^n. \end{aligned}$$

Note that

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots, \\ e^{-x} &= 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \cdots. \end{aligned}$$

Therefore,

$$\frac{e^x + e^{-x}}{2} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots,$$

and thus

$$\frac{e^{-t/\sqrt{n}} + e^{t/\sqrt{n}}}{2} = 1 + \frac{t^2}{2n} + \frac{t^4}{4! \cdot n^2} + \cdots.$$

It follows that

$$\begin{aligned} M_n(t) &= \big(1 + \frac{t^2}{2n} + \frac{t^4}{4! \cdot n^2} + \cdots\big)^n \\ &= \big(1 + (\frac{t^2}{2n} + \frac{t^4}{4! \cdot n^2} + \cdots)\big)^{\left(\frac{t^2}{2n} + \frac{t^4}{4! \cdot n^2} + \cdots\right)^{-1} \times n \times \left(\frac{t^2}{2n} + \frac{t^4}{4! \cdot n^2} + \cdots\right)} \\ &\to e^{t^2/2}, \quad \text{as } n \to \infty. \end{aligned}$$

Since $e^{t^2/2}$ is the mgf of the standard normal distribution, we conclude that

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{V[S_n]}} \xrightarrow{d} N(0,1)$$

as $n \to \infty$.

Problem 4. Each new book donated to a library needs to be processed by the librarian. Suppose that the time it takes to process a book is a random variable with mean 9 minutes and standard deviation 4 minutes.

(i) Suppose that the librarian has 50 books to process. Approximate the probability that it will take more than 480 minutes to process all the books.
(ii) Approximate the probability that at least 10 books will be processed in the first 80 minutes.
(iii) When solving the above two parts, what assumptions have you made implicitly?

Solution. (i) Let X_i $(i \geqslant 1)$ denote the time it takes to process the i-th book. We assume that this sequence of random variables are independent and identically distributed.. The amount of time it takes to process the 50 books is given by

$$S_{50} = X_1 + \cdots + X_{50}.$$

By the central limit theorem,

$$\begin{aligned}\mathbb{P}(S_{50} > 480) &= \mathbb{P}\big(\frac{S_{50} - \mathbb{E}[S_{50}]}{\sqrt{V[S_{50}]}} > \frac{480 - 50 \times 9}{\sqrt{50} \times 4}\big) \\ &\approx 1 - \Phi(1.06) \approx 0.14457.\end{aligned}$$

(ii) The desired probability is given by

$$\begin{aligned}\mathbb{P}(S_{10} \leqslant 80) &= \mathbb{P}\big(\frac{S_{10} - \mathbb{E}[S_{10}]}{\sqrt{V[S_{10}]}} \leqslant \frac{80 - 10 \times 9}{\sqrt{10} \times 4}\big) \\ &\approx \Phi(-0.79) = 1 - \Phi(0.79) \\ &\approx 0.21476.\end{aligned}$$

(iii) The assumptions we are implicitly making is that $X_1, X_2, X_3, \cdots$ are *independent* and *identically distributed.*

Problem 5. *In this problem, you may need to use the following so-called inclusion-exclusion principle without proof. Let* $A_1, A_2, \cdots, A_n$ *be* n *events. Then*

$$\begin{aligned}\mathbb{P}(A_1 \cup \cdots \cup A_n) = &\sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{1 \leqslant i < j \leqslant n} \mathbb{P}(A_i \cap A_j) \\ &+ \sum_{1 \leqslant i < j < k \leqslant n} \mathbb{P}(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1}\mathbb{P}(A_1 \cap \cdots \cap A_n).\end{aligned}$$

A little girl is painting on a blank paper. Suppose that there is a total number of N available colours. At each time she selects one colour randomly and paints on the paper. It is possible that she picks a colour that she has already used before. Different selections are assumed to be independent.

(1) Suppose that the little girl makes n selections.
(1-i) If red and blue are among the available colours, let R (respectively, B) be the event that her painting contains colour red (respectively, blue). What is $\mathbb{P}(R)$ and $\mathbb{P}(R \cup B)$?
(1-ii) Suppose that $n = N$. For $1 \leqslant i \leqslant N$, let E_i be the probability that her painting does not contain colour i. By applying the inclusion-exclusion principle to the event $\cup_{i=1}^{N} E_i$, show that

$$N! = \sum_{k=0}^{N} (-1)^k \binom{N}{k} (N-k)^N.$$

(1-iii) Let D be the number of different colours she obtain among her n selections. By writing $N - D$ as a sum of Bernoulli random variables, compute $\mathbb{E}[D]$ and $\mathrm{Var}[D]$.

(2) Let S be the number of selections needed until every available colour has been selected by the little girl. For $0 \leqslant i \leqslant N-1$, let X_i be the random variable that after obtaining i different colours, the number of extra selections needed until further receiving a new colour. By studying the distributions of these X_i's as well as their relationship with S, show that

$$\mathbb{E}[S] = N \times \big(1 + \frac{1}{2} + \cdots + \frac{1}{N}\big).$$

(3) Let T be the number of selections until the little girl picks a colour that she has obtained previously.
(3-i) By using the formula

$$\mathbb{E}[T] = \sum_{k=0}^{\infty} \mathbb{P}(T > k),$$

show that

$$\mathbb{E}[T] = \frac{N!}{N^N} \sum_{j=0}^{N} \frac{N^j}{j!}$$

(3-ii) Consider $\mathbb{E}[T]$ as a function of N. Show that

$$\mathbb{E}[T] \sim \sqrt{\frac{\pi N}{2}} \quad \text{as } N \to \infty.$$

Here the notation $a_N \sim b_N$ means $\lim_{N\to\infty} \frac{a_N}{b_N} = 1$. [*Hint: Try to use the central limit theorem in a suitable context.*]

Solution. (1-i) It is easier to work with the complements:

$$\mathbb{P}(R) = 1 - \mathbb{P}(R^c) = 1 - \big(\frac{N-1}{N}\big)^n,$$
$$\mathbb{P}(R \cup B) = 1 - \mathbb{P}(R^c \cap B^c) = 1 - \big(\frac{N-2}{N}\big)^n.$$

(1-ii) According to the inclusion-exclusion principle, we have

$$\mathbb{P}\big(\bigcup_{i=1}^N E_i\big) = \sum_{i=1}^N \mathbb{P}(E_i) - \sum_{1\leqslant i<j\leqslant N} \mathbb{P}(E_i \cap E_j) + \sum_{1\leqslant i<j<k\leqslant N} \mathbb{P}(E_i \cap E_j \cap E_k)$$
$$- \cdots + (-1)^{N-1}\mathbb{P}(E_1 \cap \cdots \cap E_N). \tag{1}$$

On the one hand, by the definition of E_i, we know that $E_1^c \cap \cdots \cap E_N^c$ is event that in the N selections the little girl obtains the complete set of colours. Its probability is given by

$$\mathbb{P}(E_1^c \cap \cdots \cap E_N^c) = \frac{N \times (N-1) \times \cdots \times 2 \times 1}{N^N} = \frac{N!}{N^N}.$$

Therefore,

$$\mathbb{P}\big(\bigcup_{i=1}^N E_i\big) = 1 - \frac{N!}{N^N}.$$

On the other hand, we have

$$\mathbb{P}(E_i) = \big(\frac{N-1}{N}\big)^N,\ \mathbb{P}(E_i \cap E_j) = \big(\frac{N-2}{N}\big)^N,\ 1 \leqslant i < j \leqslant N,$$

and more generally,

$$\mathbb{P}(E_{i_1} \cap \cdots \cap E_{i_k}) = \big(\frac{N-k}{N}\big)^N,\ 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant N.$$

By substituting these probabilities into the relation (1), we obtain

$$
\begin{aligned}
1-\frac{N!}{N^N} &= \sum_{i=1}^{N}\big(\frac{N-1}{N}\big)^N - \sum_{1\leqslant i<j\leqslant N}\big(\frac{N-2}{N}\big)^N + \sum_{1\leqslant i<j<k\leqslant N}\big(\frac{N-3}{N}\big)^N \\
&\quad - \cdots + (-1)^{N-1}\big(\frac{N-N}{N}\big)^N \\
&= \sum_{k=1}^{N}(-1)^{k-1}\binom{N}{k}\big(\frac{N-k}{N}\big)^N,
\end{aligned}
$$

which simplifies to

$$
N! = \sum_{k=0}^{N}(-1)^k\binom{N}{k}(N-k)^N.
$$

(1-iii) Let us define $Y = N - D$. By the definition of D, the random variable Y counts the number of colours that the little girl does not obtain among her n selections. We are going to express Y as a sum of Bernoulli random variables (indicator random variables) in order to calculate $\mathbb{E}[Y]$ and $V[Y]$.

For each $1 \leqslant i \leqslant N$, let E_i be the event that her selections do not contain colour i. It follows that

$$
\mathbb{P}(E_i) = \big(\frac{N-1}{N}\big)^n.
$$

We define the indicator random variable X_{E_i} by

$$
X_{E_i}(\omega) = \begin{cases} 1, & \omega \in E_i, \\ 0, & \text{otherwise.} \end{cases}
$$

Note that each X_{E_i} is a Bernoulli random variable with parameter $\mathbb{P}(E_i)$. Since Y counts the number of colours that are not contained in her selections, we have

$$
Y = X_{E_1} + \cdots + X_{E_N}.
$$

Therefore,

$$
\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}[X_{E_1}] + \cdots + \mathbb{E}[X_{E_N}] \\
&= \mathbb{P}(E_1) + \cdots + \mathbb{P}(E_N) \\
&= N \times \big(\frac{N-1}{N}\big)^n.
\end{aligned}
$$

It follows that

$$\mathbb{E}[D] = N - \mathbb{E}[Y] = N - N \times \left(\frac{N-1}{N}\right)^n.$$

To compute $V[D]$, we use the relation $V[D] = V[Y]$. Note that $X_{E_i} \cdot X_{E_j} = X_{E_i \cap E_j}$, where $X_{E_i \cap E_j}$ is the indicator random variable for the event $E_i \cap E_j$:

$$X_{E_i \cap E_j}(\omega) = \begin{cases} 1, & \omega \in E_i \cap E_j, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, we have

$$\begin{aligned} Y^2 &= \sum_{i=1}^{N} X_{E_i}^2 + 2 \sum_{1 \leqslant i < j \leqslant N} X_{E_i} \cdot X_{E_j} \\ &= \sum_{i=1}^{N} X_{E_i} + 2 \sum_{1 \leqslant i < j \leqslant N} X_{E_i \cap E_j}. \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{E}[Y^2] &= \sum_{i=1}^{N} \mathbb{E}[X_{E_i}] + 2 \sum_{1 \leqslant i < j \leqslant N} \mathbb{E}[X_{E_i \cap E_j}] \\ &= \sum_{i=1}^{N} \mathbb{P}(E_i) + 2 \sum_{1 \leqslant i < j \leqslant N} \mathbb{P}(E_i \cap E_j). \\ &= N \times (\frac{N-1}{N})^n + 2 \times \binom{N}{2} (\frac{N-2}{N})^n. \end{aligned}$$

As we result, we obtain

$$\begin{aligned} V[D] = V[Y] &= \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\ &= N \times (\frac{N-1}{N})^n + 2 \times \binom{N}{2} (\frac{N-2}{N})^n - N^2 \times (\frac{N-1}{N})^{2n}. \end{aligned}$$

Remark. When N is fixed and we view $\mathbb{E}[D], V[D]$ as functions of n, we see that $\mathbb{E}[D] \to N$ and $V[D] \to 0$ as $n \to \infty$. In particular, this suggests that, when the number n of selections is very large, $D \approx N$. Therefore, the little girl will obtain all the different colours eventually.

(2) The set of possible values of S is $\{N, N+1, N+2, \cdots\}$. By definition we have $X_0 = 1$. In addition, for each $1 \leqslant i \leqslant N-1$, after obtaining i different colours, if

we interpret obtaining an additional new colour as success, then $X_i - 1$ (number of failures before first success) is a geometric random variable with $\mathbb{P}(\text{success}) = \frac{N-i}{N}$. Therefore,

$$\mathbb{E}[X_i] = 1 + \frac{1 - \frac{N-i}{N}}{\frac{N-i}{N}} = \frac{N}{N-i}.$$

Now the key observation is that

$$S = X_0 + X_1 + \cdots + X_{N-1}.$$

Therefore,

$$\begin{aligned}
\mathbb{E}[S] &= \mathbb{E}[X_0] + \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_{N-1}] \\
&= 1 + \frac{N}{N-1} + \cdots + \frac{N}{1} \\
&= N \times \big(1 + \frac{1}{2} + \cdots + \frac{1}{N}\big).
\end{aligned}$$

Remark. Since the harmonic series $H(N) = 1 + \frac{1}{2} + \cdots + \frac{1}{N}$ *has logarithmic growth (i.e.* $\frac{H(N)}{\log N} \to 1$ *as* $N \to \infty$*), this result suggests that when* N *is large, on average the little girl needs to make about* $N \log N$ *selections before obtaining all the different colours.*

(3-i) The possible values of T are given by $T = 2, 3, \cdots, N+1$. For $1 \leqslant k \leqslant N+1$, we have

$$\mathbb{P}(T > k) = \frac{N \cdot (N-1) \cdots (N-k+1)}{N^k}.$$

Note that the above formula includes the extreme cases

$$\mathbb{P}(T > 1) = 1, \ \mathbb{P}(T > N+1) = 0.$$

Using the formula,

$$\mathbb{E}[T] = \sum_{k=0}^{\infty} \mathbb{P}(T > k),$$

we obtain that

$$\begin{aligned}
\mathbb{E}[T] &= 1 + \sum_{k=1}^{N} \frac{N \cdot (N-1) \cdots (N-k+1)}{N^k} \\
&= \sum_{k=0}^{N} \frac{N!}{(N-k)! N^k} \\
&= \sum_{j=0}^{N} \frac{N!}{j! N^{N-j}} \quad (\text{ change index to } j = N-k) \\
&= \frac{N!}{N^N} \sum_{j=0}^{N} \frac{N^j}{j!}. \qquad (2)
\end{aligned}$$

(3-ii) We need to find the asymptotic behaviour of (2) as $N \to \infty$. Let us write

$$\mathbb{E}[T] = \frac{N!}{N^N} e^N \times e^{-N} \sum_{j=0}^{N} \frac{N^j}{j!}.$$

Here the key observation is that,

$$e^{-N} \sum_{j=0}^{N} \frac{N^j}{j!} = \mathbb{P}(Z \leqslant N),$$

where Z is a Poisson random variable with parameter N. Since

$$X \sim \mathrm{Pn}(\lambda), Y \sim \mathrm{Pn}(\mu), \ X, Y \text{ independent} \implies X + Y \sim \mathrm{Pn}(\lambda + \mu),$$

we can equivalently view Z as the sum of N independent Poisson random variables with parameter 1:

$$Z = X_1 + \cdots + X_N.$$

According to the central limit theorem, we know that

$$\frac{Z - \mathbb{E}[Z]}{\sqrt{V[Z]}} \xrightarrow{d} N(0,1) \quad \text{as } N \to \infty.$$

In particular,

$$\lim_{N \to \infty} \mathbb{P}(Z \leqslant N) = \lim_{N \to \infty} \mathbb{P}\Big(\frac{Z - N}{\sqrt{N}} \leqslant 0\Big) = \Phi(0) = \frac{1}{2}.$$

In other words, we have

$$\lim_{N\to\infty} e^{-N}\sum_{j=0}^{N}\frac{N^j}{j!}=\frac{1}{2}.$$

Therefore, as $N\to\infty$, the quantity $\mathbb{E}[T]$ behaves like

$$\mathbb{E}[T]\sim\frac{1}{2}\cdot\frac{N!}{N^N}e^N\quad\text{as } N\to\infty.$$

On the other hand, from Stirling's formula for the factorial function $N!$, we know that

$$N!\sim\sqrt{2\pi N}\big(\frac{N}{e}\big)^N\quad\text{as } N\to\infty.$$

Therefore, we conclude that

$$\mathbb{E}[T]\sim\sqrt{\frac{\pi N}{2}}\quad\text{as } N\to\infty.$$

The contents below is designed for a deeper level of understanding for serious mathematicians:

Although the above solution is quite elegant, the mathematical mechanism behind the central limit theorem is rather deep (the proof of the central limit theorem does not give us much insight about why it should be true heuristically). Let us give an alternative derivation simply based on calculus without using the central limit theorem.

We start by deriving an integral representation of $e^{-N}\sum_{j=0}^{N}\frac{N^j}{j!}$. More precisely, we have

$$\begin{aligned}
&\frac{1}{N!}\int_N^\infty t^N e^{-t}dt\\
&=\frac{N^N e^{-N}}{N!}+\frac{1}{(N-1)!}\int_N^\infty t^{N-1}e^{-t}dt\quad\text{(integrate by parts)}\\
&=\frac{N^N e^{-N}}{N!}+\frac{N^{N-1}e^{-N}}{(N-1)!}+\frac{1}{(N-2)!}\int_N^\infty t^{N-2}e^{-t}dt\quad\text{(integrate by parts again)}\\
&\cdots\\
&=e^{-N}\sum_{j=0}^{N}\frac{N^j}{j!}. \qquad (3)
\end{aligned}$$

Remark. Note that this yields an interesting property that

$$\mathbb{P}(X \leqslant N) = \mathbb{P}(Y > N),$$

where $X \sim \mathrm{Pn}(N)$ *and* $Y \sim \gamma(N+1, 1)$. *This fact is not surprising as seen from the following heuristic argument. Suppose that buses arrive at a rate of one per hour. Then* $X \sim \mathrm{Pn}(N)$ *represents the number of bus arrivals among the first N hours. Therefore,* $X \leqslant N$ *is equivalent to saying that the time of the* $(N+1)$*-th arrival is greater than* N. *On the other hand, we know that the waiting time between two successive arrivals follows the exponential distribution with parameter* 1. *Therefore, the time of the* $(N+1)$*-th arrival is the sum of* $N+1$ *independent exponential distributions each with parameter* 1. *This is the Gamma distribution with parameters* $N+1$ *and* 1.

Returning to our problem, from the relation (3) we see that

$$\mathbb{E}[T] = \frac{N!}{N^N} \sum_{j=0}^{N} \frac{N^j}{j!} = \frac{e^N}{N^N} \int_N^\infty t^N \mathrm{e}^{-t} dt.$$

By applying a change of variables $t = N \times (1+s)$, the integral becomes

$$\int_N^\infty t^N e^{-t} dt = N^{N+1} e^{-N} \int_0^\infty (1+s)^N e^{-Ns} ds.$$

Therefore,

$$\mathbb{E}[T] = N \int_0^\infty (1+s)^N e^{-Ns} ds = N \int_0^\infty e^{-N(s-\log(1+s))} ds.$$

We now analyse the behaviour of the integral

$$I_N = \int_0^\infty e^{-N(s-\log(1+s))} ds$$

as $N \to \infty$.

First of all, we apply a change of variables to transform the integral to something which looks like a normal density. More precisely, we set

$$\frac{u^2}{2} = s - \log(1+s). \tag{4}$$

If we denote the right hand side (4) as $f(s)$, standard calculus shows that $f(0) = 0$ and $f(s)$ is strictly increasing to ∞ as $s \to \infty$. Therefore, $s \mapsto u$ is a legal change of variables. By differentiation, we have

$$udu = ds - \frac{ds}{1+s} = \frac{sds}{1+s},$$

so that

$$ds = \frac{(1+s)u}{s} du.$$

Therefore, we have

$$I_N = \int_0^\infty e^{-Nu^2/2} \frac{(1+s)u}{s} du.$$

Note that in the above integral, s should be viewed as a function of u which is the inverse of f, i.e. $s = f^{-1}(u)$.

To analyse $\frac{(1+s)u}{s}$ as a function of u, let us begin by observing that

$$\frac{u^2}{2} = s - \log(1+s) = \frac{s^2}{2} - \frac{s^3}{3} + O(s^4) \quad \text{when } s \text{ is small.}$$

Here we have used the Taylor expansion of the function $\log(1+s)$, and the notation $O(s^4)$ denotes a function that is comparable in magnitude with s^4, i.e. there exists $M > 0$ such that

$$|O(s^4)| \leqslant M \cdot s^4 \quad \text{when } s \text{ is small.}$$

The precise shape of $O(s^4)$ is not relevant and can change from line to line. It follows that

$$\frac{u^2}{s^2} = 1 - \frac{2}{3}s + O(s^2) \quad \text{when } s \text{ is small,}$$

Therefore,

$$\frac{u}{s} = \left(1 - \frac{2}{3}s + O(s^2)\right)^{1/2} = 1 - \frac{1}{3}s + O(s^2),$$

where we have used the fact that

$$(1+x)^\alpha = 1 + \alpha x + O(x^2) \quad \text{when } x \text{ is small,}$$

which is clear from the extended binomial theorem. As a result, we obtain that

$$\frac{(1+s)u}{s} = (1+s)\left(1 - \frac{1}{3}s + O(s^2)\right) = 1 + \frac{2}{3}s + O(s^2).$$

Note that up to first order we have $u = s$. Consequently, we arrive at the following relation:

$$\frac{(1+s)u}{s} = 1 + \frac{2}{3}u + O(u^2)$$

when s is small (or equivalently, when u is small). To summarise what we have obtained so far, we can now write

$$\frac{(1+s)u}{s} = 1 + \frac{2}{3}u + g(u),$$

where $g(u)$ is a function such that

$$|g(u)| \leqslant Cu^2, \quad \text{for all } 0 < u < c_1 \tag{5}$$

with some constants $c_1, C > 0$. The values of these constants are of no importance, and the crucial point is that $g(u)$ is of order u^2 when u is small.

To proceed further, we decompose the integral I_N into two parts:

$$I_N = J_N + K_N,$$

where

$$J_N = \int_0^{c_1} e^{-Nu^2/2}\frac{(1+s)u}{s}du, \;\; K_N = \int_{c_1}^{\infty} e^{-Nu^2/2}\frac{(1+s)u}{s}du.$$

For the K_N part, since $s \mapsto u$ is strictly increasing, with some constant c_2 we know that

$$u > c_1 \iff s > c_2.$$

Therefore,

$$\begin{aligned}
K_N &= \int_{c_1}^{\infty} e^{-Nu^2/2}\big(1+\frac{1}{s}\big)udu \\
&\leqslant \big(1+\frac{1}{c_2}\big)\cdot\int_{c_1}^{\infty} e^{-Nu^2/2}udu \\
&= \big(1+\frac{1}{c_2}\big)\cdot\frac{1}{N}e^{-Nc_1^2/2}.
\end{aligned}$$

In other words, K_N has an exponential decay as $N \to \infty$. For the J_N part, we have

$$
\begin{aligned}
J_N &= \int_0^{c_1} e^{-Nu^2/2}\big(1 + \frac{2}{3}u + g(u)\big)du \\
&= \int_0^{\infty} \big(1 + \frac{2}{3}u\big)e^{-Nu^2/2}du - \int_{c_1}^{\infty} \big(1 + \frac{2}{3}u\big)e^{-Nu^2/2}du \\
&\quad + \int_0^{c_1} g(u)e^{-Nu^2/2}du.
\end{aligned}
$$

Let us denote the above three integrals as J_N^1, J_N^2, J_N^3 respectively.

For the J_N^1 integral, it can be evaluated explicitly as

$$
\begin{aligned}
J_N^1 &= \int_0^{\infty} e^{-Nu^2/2}du + \frac{2}{3}\int_0^{\infty} ue^{-Nu^2/2}du \\
&= \frac{1}{\sqrt{N}}\int_0^{\infty} e^{-v^2/2}dv + \frac{2}{3N}\int_0^{\infty} e^{-Nu^2/2}d\big(\frac{Nu^2}{2}\big) \\
&= \frac{\sqrt{2\pi}}{2\sqrt{N}}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-v^2/2}dv + \frac{2}{3N} \\
&= \sqrt{\frac{\pi}{2N}} + \frac{2}{3N}.
\end{aligned}
$$

For the J_N^2 integral, similar to the K_N case we see that $\int_{c_1}^{\infty} \frac{2}{3}ue^{-Nu^2/2}du$ has an exponential decay as $N \to \infty$, and

$$
\int_{c_1}^{\infty} e^{-Nu^2/2}du \leqslant \int_{c_1}^{\infty} e^{-Nc_1u/2}du = \frac{2}{Nc_1}e^{-Nc_1^2/2},
$$

which is also an exponential decay. As a result, J_N^2 has an exponential decay as $N \to \infty$. For the J_N^3 integral, according to (5) we have

$$
\begin{aligned}
J_N^3 &= \int_0^{c_1} g(u)e^{-Nu^2/2}du \leqslant \int_0^{c_1} Cu^2e^{-Nu^2/2}du \\
&\leqslant \int_0^{\infty} Cu^2e^{-Nu^2/2}du = C\frac{1}{N^{3/2}}\int_0^{\infty} v^2e^{-v^2/2}dv \\
&= O(N^{-3/2})
\end{aligned}
$$

when N is large.

Putting everything together, we conclude that

$$\begin{aligned}\mathbb{E}[T] &= N \cdot I_N \\ &= N \cdot \big(\sqrt{\frac{\pi}{2N}} + \frac{2}{3N} + O(N^{-3/2}) + \text{exponential decay}\big) \\ &= \sqrt{\frac{\pi N}{2}} + \frac{2}{3} + O(N^{-1/2}) + \text{exponential decay}.\end{aligned}$$

In particular, we have

$$\mathbb{E}[T] \sim \sqrt{\pi N/2} \quad \text{as } N \to \infty.$$

This is consistent with the result we have obtained by using the central limit theorem. The extra benefit here is that we also know the magnitude of the error:

$$\frac{\mathbb{E}[T]}{\sqrt{\pi N/2}} - 1 = \frac{2\sqrt{2}}{3\sqrt{\pi}} \cdot \frac{1}{\sqrt{N}} + O(N^{-1}).$$

Therefore, the different between $\frac{\mathbb{E}[T]}{\sqrt{\pi N/2}}$ and 1 is of order $\frac{1}{\sqrt{N}}$.