



北京航空航天大学
BEIHANG UNIVERSITY

数理统计大作业

对 5G 时代下短视频业务平均请求量的统计

学 院 名 称 计算机学院

专 业 名 称 计算机科学与技术

学 生 姓 名 彭冠涵

学 生 学 号 BY2206005

课 程 教 师 孙海燕

2022 年 10 月



摘 要

在 5G 时代下，随着用户设备的增多，用户业务请求的流量增加，核心网带宽已经越来越难以承受较高的负载，需要进行边缘缓存以减少核心网流量负载。针对 5G 时代的业务流量种类特点以及新型应用的访问频率趋势，选择典型的短视频应用产生的流量进行统计量的分析，以方便未来研究边缘缓存在大量请求命中边缘时，为核心网节省的流量。Monte Carlo 算法作为计算机模拟统计中常用的算法，也可以很好地应用在计算用户请求短视频的统计量特征分析中。

本次研究对用户访问短视频产生流量的过程进行建模，使用平均请求数据下载量作为统计量去研究其分布以及具有的特征。本次研究先使用理论推导的方式得出统计量的分布，从而得出可以分析出其满足的新的参数，将其转化为标准正态分布，得出分位点函数与累计概率函数。另外，本次研究使用 Monte Carlo 算法进行请求生成，并根据其后续的概率进行模拟，得出一次模拟实验的统计量。针对分位点函数与累计概率函数的每一个值，将其转化为事件，对所有已得出统计量满足事件的频率进行统计，从而将频率作为该函数的值，并与理论计算出的函数值进行对比。本次研究还给出了 Monte Carlo 算法的误差关于实验次数的分析。

最后，对实验结果的特征进行了总结，比较了蒙特卡罗方法与其它算法的优缺点，并对 Monte Carlo 算法之后的应用进行了展望。

关键词：Monte Carlo 算法，短视频，统计量，分位函数，累计概率函数



目 录

1 引言	1
1.1 研究背景	1
1.1.1 网络用户数量与流量激增	1
1.1.2 短视频业务兴起	1
1.1.3 短视频业务特征	2
1.1.4 Monte Carlo 算法	2
1.2 研究目标与内容	2
1.3 论文组织结构	3
2 研究内容	4
2.1 问题描述	4
2.2 统计量描述	5
3 累计概率与分位点计算	6
3.1 统计量分布分析	6
3.2 参数数值	6
3.3 统计量的概率与分位点计算	7
3.4 Monte Carlo 方法概率与分位点计算	9
3.4.1 数据生成	9
3.4.2 累计概率计算	9
3.4.3 分位点函数计算	10
3.5 Monte Carlo 误差分析	11
4 总结	14
4.1 实验结果总结	14
4.2 蒙特卡洛方法与其他方法的比较	14
4.2.1 Monte Carlo 思想	14
4.2.2 Monte Carlo 方法优点	14
4.2.3 Monte Carlo 方法缺点	14
4.3 对 Monte Carlo 方法的展望	15
参考文献	16



1 引言

1.1 研究背景

1.1.1 网络用户数量与流量激增

随着 5G 时代的发展，连接到互联网的终端也处于快速增加的趋势，网络入户数量以及人均持有的联网设备增多。根据思科预测报告^[1]显示：2023 年全球人口将达到 80 亿。全球将有 53 亿人使用互联网，占比 66%，年复合增长率为 6%，人均联网设备 3.6 台，共 293 亿台，年复合增长率 10%。物联网相关的机器与机器间的连接为 147 亿，占所有连接中的 50%，年复合增长率 30%，平均移动连接速度 43.9Mbps，其中 14 亿台采用 5G 连接，速度可达 575Mbps。将上述百分比应用于中国，到 2025 年实时数据增长将超过 34 倍，占全球实时数据的 28%。

因此，随着网络用户的增多，互联网中的流量也呈现出爆炸性增长的趋势^[3]，带来了核心骨干路由器较大的压力。为了保证用户业务需求的前提下，降低互联网以及数据中心的网络带宽负载，考虑当前互联网中流量的特征，并针对性地使用边缘服务器对部分数据进行缓存成为了当前主流的解决方案。

1.1.2 短视频业务兴起

在 5G 时代中，随着人均拥有联网设备的上升，人们越来越习惯于使用移动设备进行平常业务的需求，短视频业务凭借着内容贴近生活容易引发情感共鸣、围绕热点话题与网友同频共振、清晰人设标签与推荐算法智能化等特征逐渐呈现爆炸性的增长，其在互联网中产生的流量已经成为了不可忽视的一部分。

事实上，根据 Cisco 报告^[3]显示，虽然 2022 年以后网络流量中物联网流量增长最为迅速，但是视频流量仍然会在较长时间内占据较大的比重。另外由于短视频业务^[6]的特征，当用户获取的静态视频缓存在边缘服务器时，可以极大地节省用户本次请求产生的流量。因此对短视频的业务进行单独的研究，分析已有的用户请求与视频特征的相关性，从而推断出未来用户访问视频规律^[7]。最终结合不同时段用户访问的规律以及边缘服务器约束，利用优化算法随时调整不同边缘服务器，从而更好地将用户请求在边缘处解决，节省核心网络带宽^[8]。



1.1.3 短视频业务特征

研究中^[9]指出,短视频业务,例如以抖音、快手^[10]等短视频服务平台的点击量显现出了与以往视频服务平台不一样的特征。由于短视频应用中实时热度榜的存在,智能推荐算法往往会将学习用户最近访问的视频特征并推荐相关标签的同类视频,因此用户的请求表现出了短时段内集中在前几位实时热点视频中,随着视频热度降低,用户请求急剧减少^[11]。在不同时段间热度榜会有调整,但是用户请求大量集中在热度较高视频趋势不会变。根据研究中给出的统计数据,现有的用户请求频率与视频热度排名关系服从帕累托规则,即一小部分视频占浏览量的很大比例。根据 Zhang 等人^[12]提出其满足 Zipf 分布^[2],最终统计出短视频业务的大致参数,后续会使用此参数进行统计。

1.1.4 Monte Carlo 算法

Monte Carlo 算法也叫随机抽样或统计试验方法^[4],其来源是为了解决核武器实验中的裂变物质扩散问题而提出的^[5]。在物理实验中,尤其是前沿理论实验中,往往真实的世界参数较为复杂,并没有很好的理论支撑,因此只能采用实验的方法得出实验值,从而得出一个经验的统计规律。而后随着计算机的出现,可以通过生成随机数以及数值模拟的计算方式进行实验,由于计算机的机械特点,可以节省真实的实验资源,从而 Monte Carlo 算法继续发展^[15]。现在的计算机已经能实现高效计算,并且随着物理问题越来越复杂,试验成本越来越高,计算机模拟的 Monte Carlo 方法已经逐渐成为主流的解决方案^[16]。

1.2 研究目标与内容

本次研究的主要内容分为三方面。第一是对用户获取短视频产生流量的过程以及用户请求分布特征进行建模分析,得到用户平均请求到的带宽在本模型下的表达式,并将其作为统计量 $T(x)$,对其分布进行分析。

第二,根据数理统计相关知识,上述的 $T(x)$ 的累计概率 α 与分位点 t 是一一对应的函数关系。第一个累积分布函数 $\alpha(t)$ 是在给定 t 的情况下,计算 $\alpha = P\{T(x_1, x_2, \dots, x_n) \leq t\}$ 的值。第二个分位点函数 $t(\alpha)$,是在给定 α 的情况下,计算 $P\{T(x_1, x_2, \dots, x_n) \leq t\} = \alpha$ 时 t 的值。本次研究使用理论计算与 Monte Carlo 的方式得到上述两个函数的结果。

第三是对于 Monte Carlo 得出的结果关于样本容量的不同进行误差分析,并对 Monte Carlo 的计算方式进行优缺点分析。



1.3 论文组织结构

第一章为引言，主要介绍了本次研究的背景。在引言的第二部分提出了本次的研究目标与内容，明确了本次对统计量进行累积分布函数计算与分位函数计算的目标，并简要叙述了本文的内容组织结构。

第二章是论文主体的研究内容介绍。在第二章中首先介绍了本次研究问题的模型，用关键变量将用户获取短视频产生流量的。在第二章的第二部分给出了本次用户产生流量的关键指标，并将其作为统计量。

第三章首先对第二节定义的统计量进行了理论分析，给出了该统计量是正态分布的。根据正态分布的性质，可以将统计量累计分布函数中每一个点的计算转为标准正态分布变量与常数的比较形式，利用计算机查找累计分布函数，得到概率。同理，将其转为标准正态分布后使用计算机进行标准正态分布的计算即可得出在给定累计分布值下分位点的函数值。另外，在第三章给出了使用 Monte Carlo 算法计算累积分布函数以及分为函数的值。

第四章主要对上述两种方法得到累计分布函数与分位点函数的结果进行分析，并给出了函数结果图。另外，讨论了 Monte Carlo 方法计算结果随着样本容量变化时计算误差的变化，并进行总结。

最后一章对本次研究进行讨论，指出了 Monte Carlo 算法在本次研究中发挥的作用，以及 Monte Carlo 算法相较于其他概率计算方法的优缺点比较分析。

2 研究内容

2.1 问题描述

本次研究内容的第一部分是对 5G 时代下移动用户观看短视频产生的流量进行建模，并根据关键指标建立统计量。

在第一章的研究背景中指出了，当前核心网最主要的瓶颈是移动用户的增多，在请求视频等数据量庞大的任务时对带宽的占用。因此关键性的指标是用户平均请求业务的数据大小。当要部署边缘服务器进行缓存以节省流量时，可以根据附近基站的用户请求历史数据进行节省带宽的估计，即其缓存高热度短视频时在用户命中的时候发给用户的数据量^[14]。最终使用用户对视频的平均请求下载量来表示对核心网的负载占用情况，将其作为统计量。

表 2.1 变量含义表描述了在建立模型的过程中涉及到的变量符号及其含义。

表 2.1 变量含义表

变量名称	对应含义
n	总请求个数
i	请求对应下标
x_i	某次请求对应的下载量
S	考虑的短视频总个数
s	短视频编号
f_s	所有请求对应到第 s 个短视频的概率
R_s	第 s 个短视频的热度排名
α	Zipf 分布的流行度参数
C	Zipf 分布中的常数
d_s	节点 i 到节点 j 的链路开销权重
(μ, σ^2)	每个短视频下载量服从的正态分布参数
(μ_T, σ_T^2)	统计量 $T(x)$ 服从的正态分布参数
β	Zipf 分布的加权下正态分布的标准差之比
m	Monte Carlo 算法实验次数

在本模型中，样本为用户发出的请求的下载量 x_i ，总数为 n 。短视频总共有 S 个，每个短视频下载的数据量为 d_s ，每个 d_s 遵从参数为 (μ, σ^2) 的正态分布。每个请求以 Zipf 的概率分布到各个短视频上，每个短视频对应的请求数量应为 $n * f_s$ 。可以知道，总请求数

被分配到每一个短视频后, 应该保持总量不变, 即分配到不同短视频的频率之和应该为 1, 如(2.1)所示。

$$\sum_s f_s = 1 \quad (2.1)$$

另外, 由 Zipf 分布可知短视频排名与请求到该短视频的概率关系如下, 其中 α 为 Zipf 分布的流行度参数, 数值越大表示排名高的短视频收到的请求越集中:

$$f_s * R_s^\alpha = C \quad (2.2)$$

由于某一时刻的排行榜已知, 统计时可以固定从当前热度高的短视频开始统计, 即固定当前短视频下标为每个短视频的 R_s , 则有:

$$R_s = s \quad (2.3)$$

再由(2.1)式和(2.2)可知:

$$C * \sum_s R_s^{-\alpha} = C * \sum_s s^{-\alpha} = 1 \quad (2.4)$$

$$f_s = \frac{s^{-\alpha}}{\sum_s s^{-\alpha}} \quad (2.5)$$

2.2 统计量描述

本次提出的统计量 $T(x_1, x_2, \dots, x_n) = \bar{x} = \frac{\sum_i x_i}{n}$, 表示用户以 Zipf 分布^[2]来请求短视频文件时, 所产生的平均请求开销, 以下简称其为 $T(x)$ 。

3 累计概率与分位点计算

3.1 统计量分布分析

该统计量中的每个 x_i 的分布不好确定，因为其所对应的短视频未知，因此我们需要按照每个短视频的下载量进行计算，最终求和即可。总流量即为每个短视频收到的请求数 $n * f_s$ 乘该短视频的数据量 d_s ，可得出平均请求量为统计量 $T(x)$ 如(3.1)所示。

$$T(x) = \sum_s \frac{n * f_s * d_s}{n} = \sum_s f_s * d_s = \sum_s \frac{s^{-\alpha} * d_s}{\sum_s s^{-\alpha}} \quad (3.1)$$

由于 $s^{-\alpha}$ 为常数，每个短视频大小 d_s 是独立同分布的，且 $d_s \sim N(\mu, \sigma^2)$ ，因此 $T(x)$ 可以表示为正态分布的线性和，因此 $T(x)$ 也为正态分布，设 $T(x) \sim N(\mu_T, \sigma_T^2)$ ，则有

$$\mu_T = \sum_s \frac{s^{-\alpha} * \mu}{\sum_s s^{-\alpha}} = \sum_s \frac{s^{-\alpha} * \mu}{\sum_s s^{-\alpha}} = \mu \frac{\sum_s s^{-\alpha}}{\sum_s s^{-\alpha}} = \mu \quad (3.2)$$

$$\sigma_T^2 = \sum_s \frac{s^{-2\alpha} * \sigma^2}{(\sum_s s^{-\alpha})^2} = \sigma^2 \sum_s \frac{s^{-2\alpha}}{(\sum_s s^{-\alpha})^2} \quad (3.3)$$

设 $\beta^2 = \sum_s \frac{s^{-2\alpha}}{(\sum_s s^{-\alpha})^2}$ ，则有：

$$\sigma_T^2 = \beta^2 \sigma^2 \quad (3.4)$$

容易知道， β 只与短视频的个数 S 有关，在考虑不同个数的短视频排行榜时，个数越多， β 越小，此处不给出证明，后续也可以通过图示发现这一特点。

3.2 参数数值

在本模型中参数较多，在此给出基于已有的统计结果的一些参数设置，后续，如表 3.1 参数设置表所示。

表 3.1 参数设置表

参数	数值
α	0.62
S	60
μ	30
σ^2	100
n	1000

后续的基于以上参数进行实验，可变的只有实验的次数，以方便后续 Monte Carlo 算法的模拟与误差分析。在理论计算分析中会将短视频 S 设置为一组数值，以更好地说明统计量的分布，但在后续 Monte Carlo 模拟过程中 S 始终为 60。

3.3 统计量的概率与分位点计算

按照统计量分布分析, $T(x) \sim N(\mu_T, \sigma_T^2)$, 其中 $\mu_T = \mu$, $\sigma_T = \beta\sigma$, 容易知道 $T(x)$ 的分布与样本大小无关。若设 $z = \frac{T(x) - \mu}{\sigma_T} \sim N(0, 1)$, z 为标准正态分布容易得出其概率密度函数、分位点函数、累计概率函数。因此对 $T(x)$ 也可以通过简单的转化为标准正态分布从而求出函数值, 依照此做法, $T(x)$ 的概率密度函数如图 3.1 统计量概率密度函数图所示。可以看出, 当短视频个数 S 增加的时候, σ_T 变小, 即 β 变小。

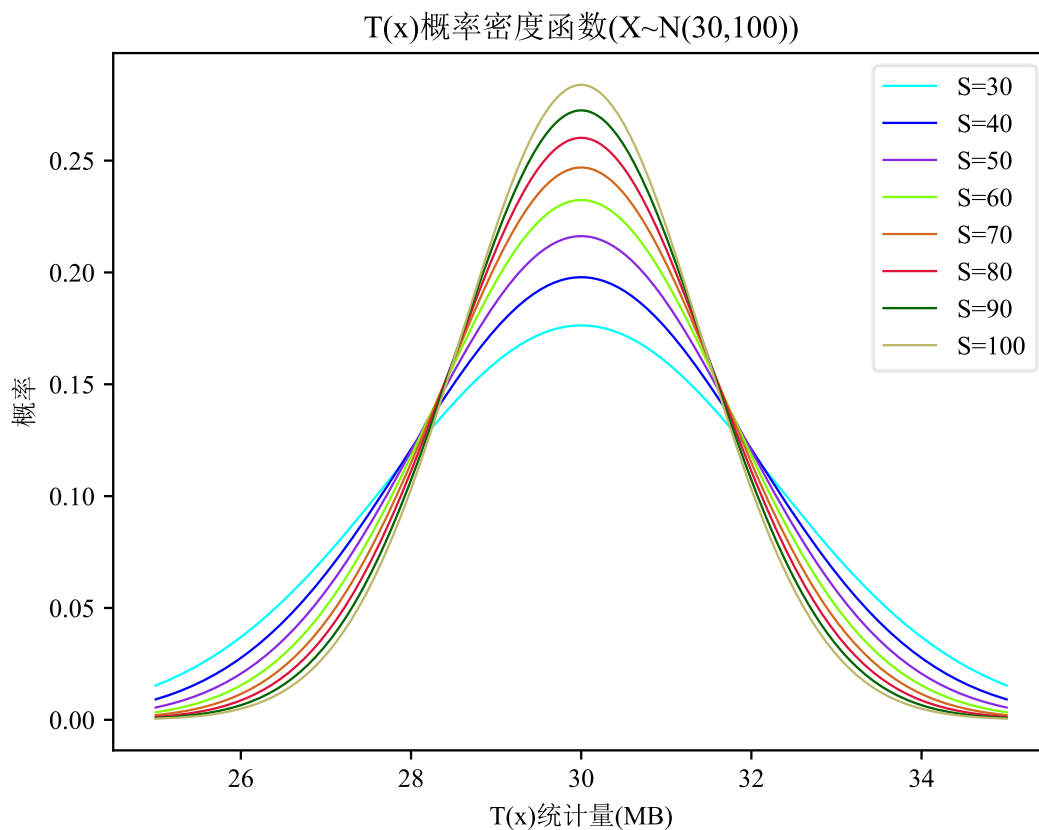
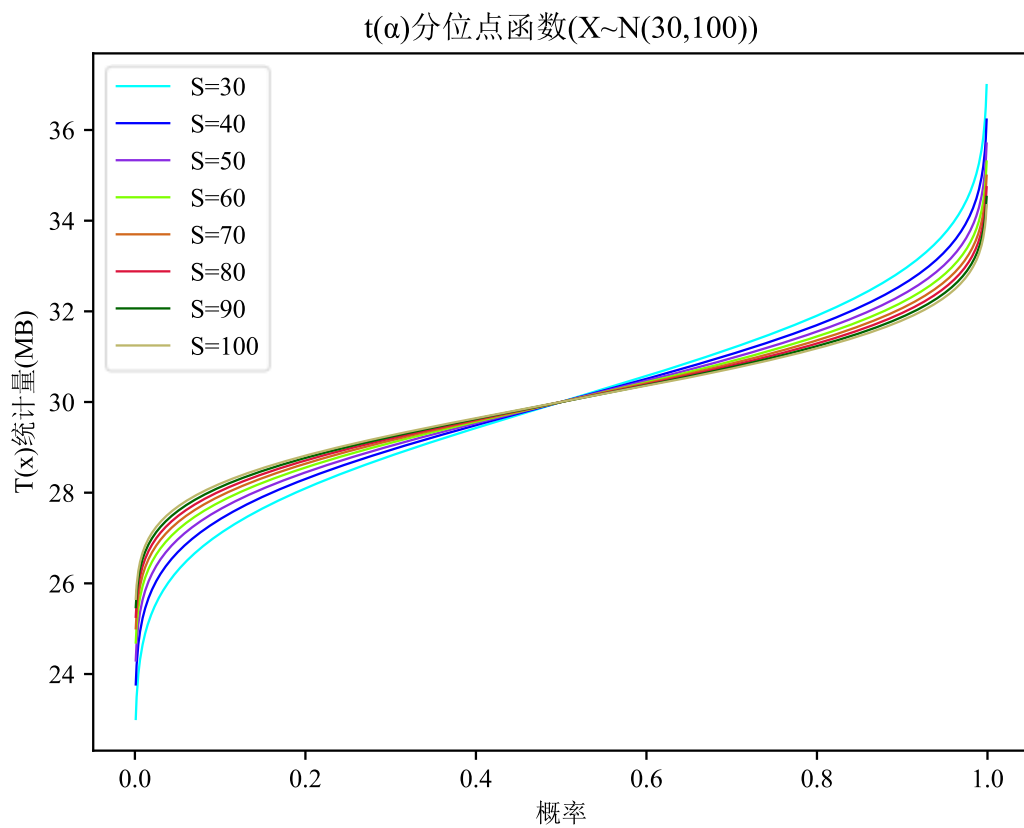
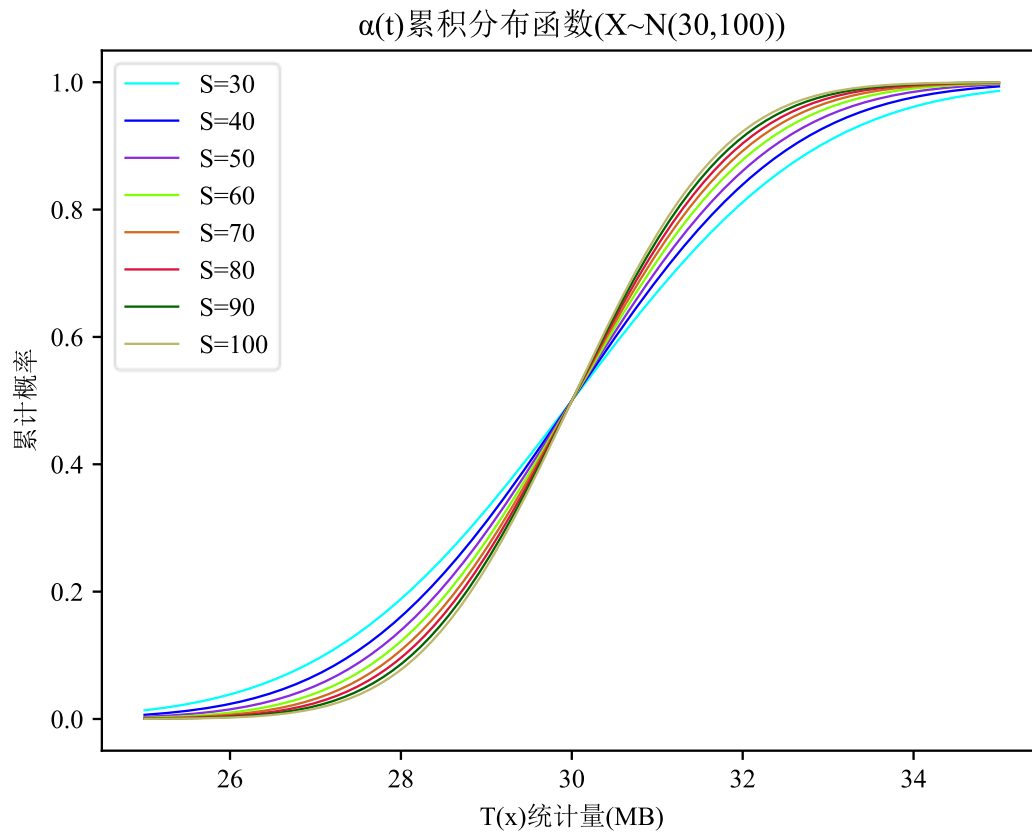


图 3.1 统计量概率密度函数图

同理, 根据现有的计算机中已经计算过的较为精确地标准正态分布函数的分位点与累计概率函数, 可以得到统计量 $T(x)$ 的累计概率函数 $\alpha(t)$ 如图 3.2 统计量累计概率函数图所示, 与分位点函数 $t(\alpha)$ 如图 3.3 统计量分位点函数图所示。

可以看出, 累计概率函数随着短视频个数的增多而变得陡峭, 分位点函数在概率接近 1 以及 0 的时候随短视频个数增多而变得陡峭。后续会将 Monte Carlo 算法得出的函数与 $S = 60$ 的函数进行比较, 并分析其准确性。





3.4 Monte Carlo 方法概率与分位点计算

使用蒙特卡洛方法来计算时，主要的思路是生成大量的请求，分别计算每个请求的下载量，再进行平均即可。但是在计算的时候涉及到请求的按概率分配，以及请求到的短视频数据需要按照正态分布生成。每一次实验得到的统计量都要记录，并最终根据统计量是否满足小于一个数值来分别计算概率，从而得到分位点与累计概率。

3.4.1 数据生成

第一步，需要按已有参数的正态分布随机生成短视频的数据大小，并进行记录。

第二步，生成 Zipf 分布的分位点函数，即将所有概率拼凑成前缀和。

第三步，生成一定数量的请求。

第四步，每个请求生成一个 0 到 1 的随机数，根据 zipf 的概率前缀和，即累计概率，选取此时请求对应到的短视频，即可得出此请求的下载量。将这些下载量进行平均即可得出统计量。

重复一到四步 m 次，分别统计出 m 个统计量，并将其记录。

3.4.2 累计概率计算

对一个上届 t ，只需要统计出当前有多少次实验的统计量小于 t ，并将其除以总实验次数 m ，得出概率，最终将该概率作为累计概率 α 。对不同的 t 进行上述计算，则可以得出累计概率函数，如图 3.4 统计量分位点函数图所示。

可以看出，在实验次数较少的时候，函数图像呈现出了锯齿状。这是因为在实验的次数比较小的时候得到的统计量个数较为离散，无法覆盖所有的值域，采用上述的累计概率函数算法在上届值不在实验结果中的时候，就会使用比上届小的最大结果作为函数值。需要注意的是，实验次数较少的时候会发现在较大 t 值下，概率并没有等于 1，这只是因为图中的结果仅仅截取了在 $[\mu - 2\sigma_T, \mu + 2\sigma_T]$ 的部分，扩大自变量范围则可以看到所有函数在 t 取较大值时函数值为 1。这也说明了实验次数较少的时候，有一些值落在距离 μ 较远的范围内会对整体曲线有较大的影响。

不仅如此，在实验次数最少的时候也可以看出每个点的函数值是和理论计算的值偏差最大的。特殊的，当实验次数为 4 是，在统计量在 29 之前的时候概率为 0，这显然是不符合预期的，正态分布应该是较为平滑的。但由于次数为 4，按照上述算法累计概率只能取 $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ ，上届落在同一个相邻概率区间内的时候函数值相同。随着实验次数增

加到 16384 时, Monte Carlo 算法累计分布函数曲线与理论计算出的函数曲线的差别几乎不可见, 已经可以有较好的结果。对模拟次数与具体的误差的分析会在之后进行介绍。

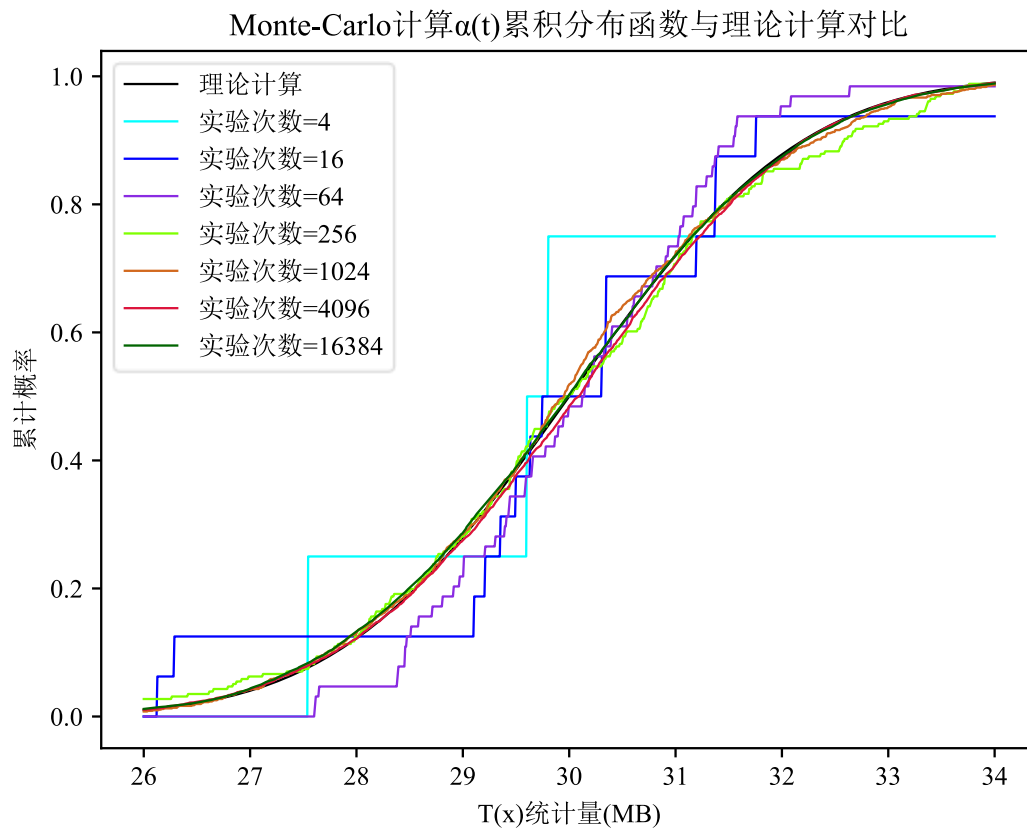


图 3.4 统计量分位点函数图

3.4.3 分位点函数计算

将 m 次实验得出的统计量按升序进行排序。从第一个到最后一个依次使用下标 j 除以 m 来代表总的累计概率 α , 由于已经按照升序排列, 则前 j 个统计量都小于第 j 个统计量, 因此前 j 个元素的上确界就是第 j 个元素, 将第 j 个元素作为 t , 从而每一个 α 都能有一个 t 与之对应, 最终可以得到分位点函数, 如图 3.5 统计量分位点函数图所示。

与累计分布函数的图类似, 分位点函数也有类似的规律。可以看出在实验次数在 64 次及以下时阶梯状的特征表现较为显著, 256 次以上的时候基本已经可以很好地体现函数特征了, 实验次数取到几千的时候就已经几乎贴合理论计算曲线, 再往上增加实验次数所减小的误差已经很难在图像中进行体现了。

值得注意的是, 分位点函数自变量取值只能为 $(0,1)$, 已经全部包含在图像中。当实验次数比较小的时候, 例如实验次数为 4 的时候, 可以看到在分位点接近 1 的时候, 函数值只略高于 30, 不到 32, 分位点接近 0 的时候比 28 小, 这显然是不对称的, 与理论

计算结果相违背。不仅如此,可以看出有三次实验的统计量大于 30,一次统计量小于 30,但是偏差最大的值刚好是那一次小于 30 的统计量。其原因就在于碰到了偶然事件,实验中有较多次数取到了 30 以上,但偏差较小,同时少量实验得到的统计量在 30 以下却因为偏差较大,拉低了下届导致了函数不对称。这也印证了在累计分布概率中所提到的实验次数较少的时候一些偏差较大的点会对整体函数特征产生较大影响。

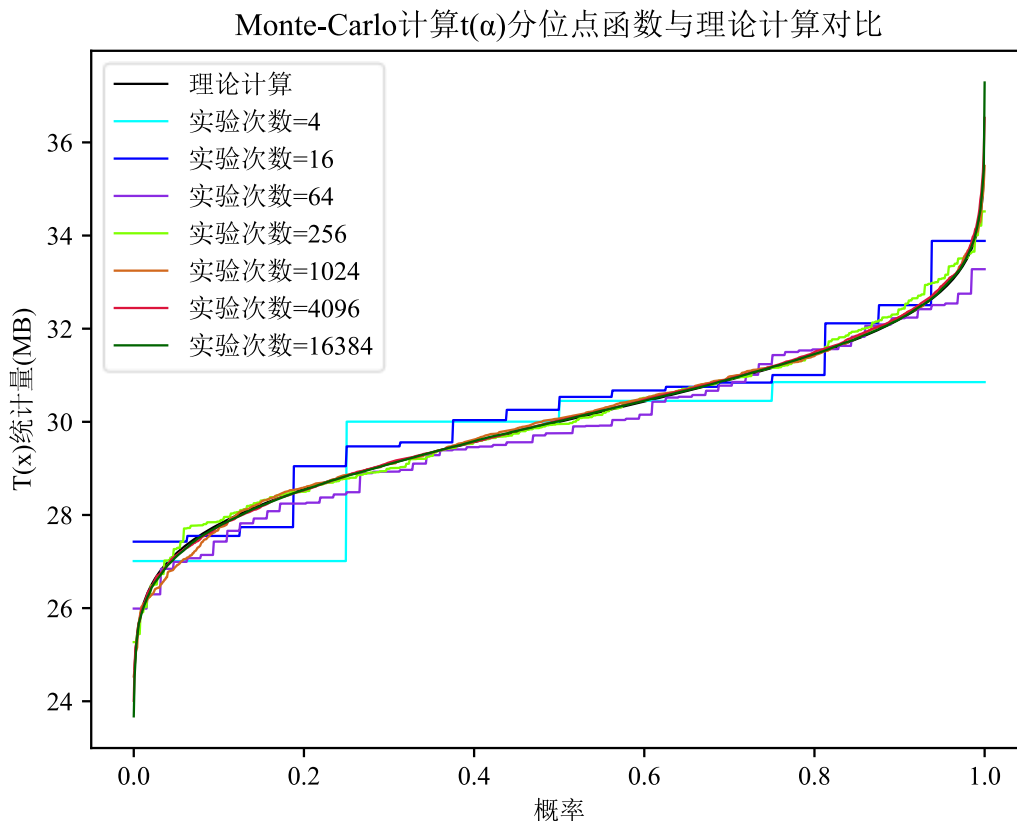


图 3.5 统计量分位点函数图

3.5 Monte Carlo 误差分析

误差从分位点函数与累计概率函数与理论计算的差距得出。在分位点函数自变量范围 0 到 1 中平均的取一些值,计算两种算法的函数值的差的绝对值,进行相对误差和绝对误差的统计。

对于相对误差,将差值的绝对值除以理论值后可以得出相对的误差占比,并对每一个点的误差比求算术平均值。对于绝对误差,采用均方误差的统计方式进行评估,将两种算法函数值的差值求平方后再进行算术平均从而得到均方误差。

由于累计概率分布函数与分位点函数互为逆函数,在探究实验次数对函数值误差的影响的时候,对其中一个函数的值进行误差分析即可代表另一个函数的准确程度。由于

分位点函数的定义域为 $(0,1)$ ，而累积分布函数的定义域为 $(-\infty, +\infty)$ ，在统计误差时需要不断取定义域上的点进行值的计算，因此使用分位点函数能够更准确方便地评估误差值。最终相对误差与绝对均方误差值随实验次数 m 的变化如图所示。

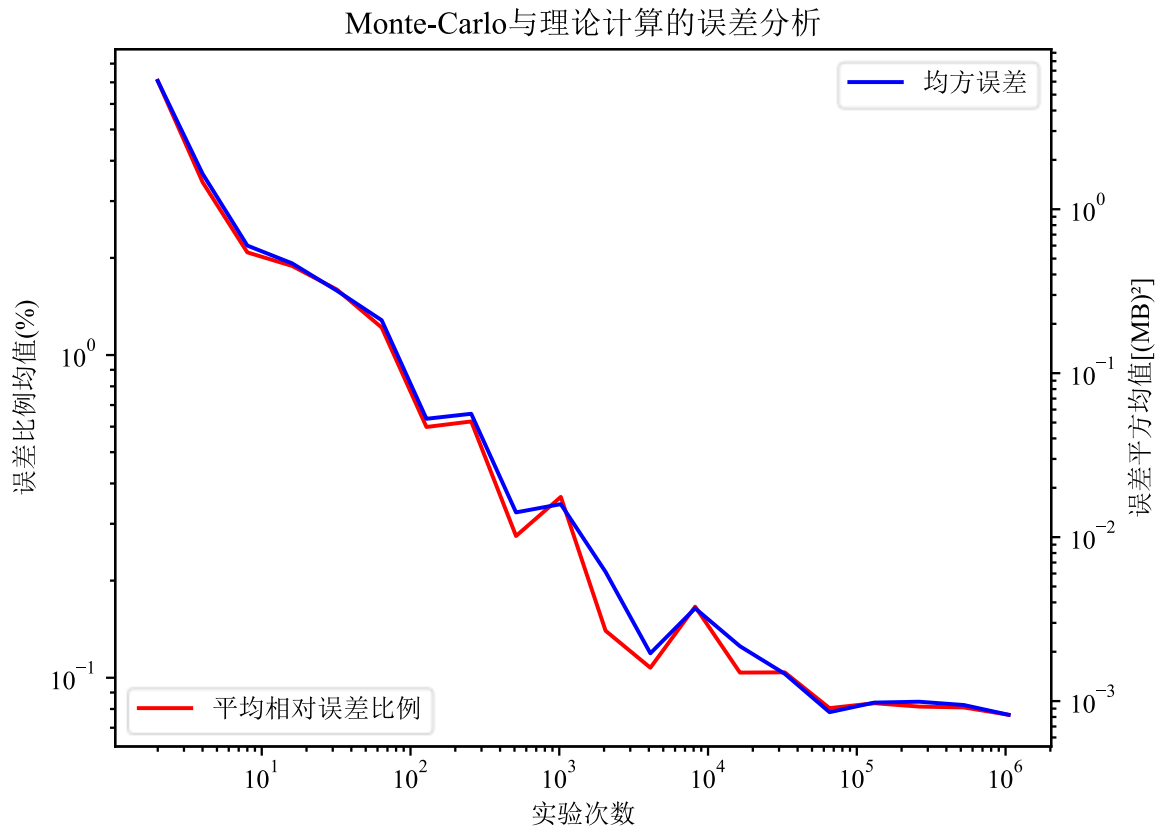


图 3.6 统计量分位点函数图

可以看到，总体来说，均方误差与平均相对误差比例随实验次数增大而下降。在图中横纵坐标均为对数分布，而图中的两种误差评价指标均随实验次数的数量级增大而线性减小数量级，即误差的对数与实验次数的对数呈线性负相关趋势，由此可以推出误差与实验次数呈反比例相关。

图中在实验次数逐渐增大的过程中，误差值可能会有抖动的情况，在局部会出现实验次数增加，但误差反而增加，这可能是由于计算机对较小数字精度的损失，以及伪随机数的生成导致的偶然情况。说明即使误差会小于一定的值，但是在当前误差取值数量级内，其取值仍然会有摆动的情況，并不固定。

当实验次数大于 10^5 时，误差值并没有继续地下降，第一方面是因为其下降的速度要变得更慢，另一方面是由于计算机精度的问题，导致。这也就说明了两个问题：第一，



并不是只要实验次数取的足够多就能绝对精确，其精确程度还会受到模拟精度、硬件设备的制约；第二，提升实验次数来减少误差存在边际效应，当实验次数足够多的时候，提升实验次数带来的精确程度会很小。

另外，由于在误差分析中，值可能跨越几个数量级，因此在表 3.2 误差数值表中给出了指标的具体值以供参考。

表 3.2 误差数值表

实验次数	平均相对误差比例(%)	均方误差(MB ²)
2	7.071496181821535	6.063014307947725
4	3.435583392486104	1.6474106217150848
8	2.080290156127427	0.6002359344575346
16	1.888688563473371	0.46771728789164896
32	1.5955484285042132	0.31872780567344383
64	1.2170398930164232	0.21041984506498582
128	0.598506869806009	0.05271512997621415
256	0.6226914173897741	0.056554057155286136
512	0.2749864090847641	0.014158057794764452
1024	0.3634064650968415	0.01589614717072891
2048	0.13981972317656577	0.006135262917031522
4096	0.10741493213481108	0.0019567026618278736
8192	0.1660774964598449	0.003681141203362325
16384	0.1037471016631555	0.002162335319540631
32768	0.10393090715963384	0.0014676381641995952
65536	0.08050216407039101	0.0008557683333915925
131072	0.08338644810027232	0.0009792830844916585
262144	0.08140377069172645	0.0009912594487388984
524288	0.08090523362656339	0.0009465022547368505
1048576	0.07676964079994042	0.0008250008039271767



4 总结

4.1 实验结果总结

第三章中的累积分布函数图像以及分位点函数图像已经能够看出，在实验次数较小的时候，Monte Carlo 算法并不能很好地表示连续的函数，但是当实验次数增多的时候，该算法绘制出的函数图像已经基本能够完全贴合预先理论计算的累计概率函数和分位点函数。

Monte Carlo 算法的误差随着实验次数数量级的增长，呈现出减少趋势，两者大致成反比例关系。同时在实验次数值较大的情况下，提升实验次数带来的精确程度收益较小。

针对本次研究问题，从数量上的统计来说，实验次数在几万次的数量级即可保证能够较为准确地反应各个函数的特征，平均相对误差比例可到千分之一。

4.2 蒙特卡洛方法与其他方法的比较

4.2.1 Monte Carlo 思想

当求解的问题或者最终的指标可以表示为一次实验中的观察值的时候，那么就可以通过模拟的方式进行实验，得到出现期待的目标事件或者某个指标值占总实验次数的比重。当做的实验次数足够多的时候，将目标事件或数值出现的概率就当做该求解的问题的概率^[13]。

4.2.2 Monte Carlo 方法优点

1. 最终结果与问题的理论推导复杂性无关，只与实验复杂度有关，即只要知道怎么模拟就可以开始生成样本，通过计算机模拟得出最终结果，经过多次实验过后对概率进行统计即可求解原问题，降低了某些问题复杂度^[15]。
2. 误差与样本规模有关，对于一些简单的重复性实验，完全可以暴力增加样本规模来降低误差，让计算机去模拟，从而较为精准地求出原问题的解。

4.2.3 Monte Carlo 方法缺点

1. 无法得出理论的精确解，即只能用作数值的计算，如果求解问题只是一个问题的子集，则由 Monte Carlo 算法不能得到理论关系式，误差可能会影响到其他问



题的求解，从而越来越放大误差。此外，无法得出精确解还可能导致其他问题只能使用数值求解，不能显式地求出各个子问题的关系，从而不能化简。

2. 模拟成本大的时候不适用，例如物理实验中需要 Monte Carlo 模拟与真实做实验相互配合检测概率，在耗费材料成本高时，则不可取。又或者在计算机进行模拟时需要较大样本收敛或者复杂度过高的时候也不能取得很好的效果^[16]。
3. 误差并不能通过对原问题的分析从而准确分析，是做实验的概率误差，不够精确。

4.3 对 Monte Carlo 方法的展望

Monte Carlo 算法比较通用，只要在求解的问题可以进行模拟或者实际操作，就可与对事件发生概率进行统计。因此随着计算机能力的提升，以及并行计算的完善，越来越多、规模越来越大的问题可以放到计算机去执行，并在可接受的时间内得出结果，采用模拟对一些比较独立而又复杂的问题，可以快速地进行数值求解，不需要对每个具体问题都通过数学推导计算出结果。因此通过对 Monte Carlo 算法的一些通用步骤分析，以及典型案例的具体操作及其带来的开销的学习，都可以更好地帮助解决现有的较复杂的问题。



参考文献

- [1] Cisco. Cisco annual internet report (2018–2023) white paper[EB/OL], <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf>, 2020-11-20/2022-2-28
- [2] G. Zipf. Human Behavior and the Principle of Least Effort. Addison-Wesley (Reading MA), 1949.
- [3] Cisco. 2021 Global Networking Trends Report[EB/OL], <https://www.cisco.com/c/dam/en/us/solutions/enterprise-networks/2021-networking-report.pdf>, 2020-11-20/2022-2-26
- [4] James F. Monte Carlo theory and practice[J]. Reports on progress in Physics, 1980, 43(9): 1145.
- [5] Metropolis N, Ulam S. The monte carlo method[J]. Journal of the American statistical association, 1949, 44(247): 335-341.
- [6] Cheng X, Dale C, Liu J. Understanding the characteristics of internet short video sharing: YouTube as a case study[J]. arXiv preprint arXiv:0707.3670, 2007.
- [7] Cheng X, Liu J, Dale C. Understanding the characteristics of internet short video sharing: A youtube-based measurement study[J]. IEEE transactions on multimedia, 2013, 15(5): 1184-1194.
- [8] Zhang Y, Li P, Zhang Z, et al. Autosight: Distributed edge caching in short video network[J]. IEEE Network, 2020, 34(3): 194-199.
- [9] Zhang Y, Li P, Zhang Z, et al. Challenges and chances for the emerging short video network[C]//IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2019: 1025-1026.
- [10] Zhang Y, Liu Y, Guo L, et al. Measurement of a Large-Scale Short-Video Service Over Mobile and Wireless Networks[J]. IEEE Transactions on Mobile Computing, 2022.
- [11] Zhang Y, Li P, Zhang Z, et al. Autosight: Distributed edge caching in short video network[J]. IEEE Network, 2020, 34(3): 194-199.
- [12] Zhang Y, Li P, Zhang Z, et al. Challenges and chances for the emerging short video network[C]//IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2019: 1025-1026.
- [13] Kalos M H, Whitlock P A. Monte carlo methods[M]. John Wiley & Sons, 2009.



-
- [14] Cha M, Kwak H, Rodriguez P, et al. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system[C]//Proceedings of the 7th ACM SIGCOMM conference on Internet measurement. 2007: 1-14.
- [15] Raychaudhuri S. Introduction to monte carlo simulation[C]//2008 Winter simulation conference. IEEE, 2008: 91-100.
- [16] Kroese D P, Brereton T, Taimre T, et al. Why the Monte Carlo method is so important today[J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2014, 6(6): 386-392.