1. 定义

假设 $X_1, X_2, \dots, X_n$ 是从一个总体中抽取的独立同分布（i.i.d.）的随机样本。将这些样本按从小到大的顺序排列，得到一组新的随机变量：
$X_{(1)} \le X_{(2)} \le \dots \le X_{(n)}$

这组有序的随机变量就称为**次序统计量**。

- $X_{(1)} = \min(X_1, X_2, \dots, X_n)$ 是样本最小值。
- $X_{(n)} = \max(X_1, X_2, \dots, X_n)$ 是样本最大值。
- $X_{(k)}$ 是第 k 小的观测值。
- 如果 n 是奇数，样本中位数是 $X_{((n+1)/2)}$。
- 样本极差是 $R = X_{(n)} - X_{(1)}$。
- 样本中程数是 $\frac{1}{2}(X_{(1)} + X_{(n)})$，就是你原题中用到的量。

2. 次序统计量的分布

这是最核心的部分。设原始样本来自一个连续分布，其累积分布函数（CDF）为 $F(x)$，概率密度函数（PDF）为 $f(x)$。

a. 第 k 个次序统计量 $X_{(k)}$ 的分布

直观理解 (推导PDF的思路):
要让第 k 小的值 $X_{(k)}$ 恰好落在很小的区间 $[x, x+dx]$ 内，必须发生三件事：

1. 有 **1** 个样本值落入 $[x, x+dx]$ 内，其概率约为 $f(x)dx$。
2. 有 **k-1** 个样本值小于 x，其概率为 $F(x)$。
3. 有 **n-k** 个样本值大于 $x+dx$（约等于大于 x），其概率为 $1-F(x)$。

这 n 个样本值可以有多种组合方式来实现上述配置。从 n 个中选 1 个落在 $[x, x+dx]$，再从剩下的 $n-1$ 个中选 $k-1$ 个小于 x，组合数为 $\binom{n}{1}\binom{n-1}{k-1} = \frac{n!}{(k-1)!(n-k)!}$。

综合起来，就得到了 $X_{(k)}$ 的概率密度函数（PDF）：

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}[1-F(x)]^{n-k}f(x)$$

特殊情况（非常常用）：

- **样本最小值 $X_{(1)}$ 的PDF ($k=1$):**
  $f_{X_{(1)}}(x) = \frac{n!}{0!(n-1)!}[F(x)]^0[1-F(x)]^{n-1}f(x) = n[1-F(x)]^{n-1}f(x)$
- **样本最大值 $X_{(n)}$ 的PDF ($k=n$):**
  $f_{X_{(n)}}(x) = \frac{n!}{(n-1)!0!}[F(x)]^{n-1}[1-F(x)]^0 f(x) = n[F(x)]^{n-1}f(x)$

b. $X_{(1)}$ 和 $X_{(n)}$ 的累积分布函数（CDF）

有时候 CDF 更容易推导和使用。

- **最大值 $X_{(n)}$ 的CDF**:
  事件 "$X_{(n)} \le x$" 等价于事件 "所有样本值都 $\le x$"。
  $F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P(X_1 \le x, \dots, X_n \le x) = [P(X \le x)]^n = [F(x)]^n$
- **最小值 $X_{(1)}$ 的CDF**:
  直接求 $P(X_{(1)} \le x)$ 不方便，我们求其对立事件。
  $F_{X_{(1)}}(x) = P(X_{(1)} \le x) = 1 - P(X_{(1)} > x)$
  事件 "$X_{(1)} > x$" 等价于事件 "所有样本值都 $> x$"。
  $P(X_{(1)} > x) = P(X_1 > x, \dots, X_n > x) = [P(X > x)]^n = [1 - F(x)]^n$
  所以，$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n$

3. 最重要的特例：标准均匀分布 $U(0,1)$

当原始样本 $X_1, \dots, X_n$ 来自 $U(0,1)$ 分布时，次序统计量的性质有非常简洁和优美的形式。这是因为对于 $x \in (0,1)$，我们有 $F(x) = x$ 且 $f(x) = 1$。

将此代入 $X_{(k)}$ 的 PDF 公式:
$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} x^{k-1}(1-x)^{n-k}$
这个形式你可能很眼熟，它就是 **Beta 分布** 的 PDF！具体来说:

$$X_{(k)} \sim \text{Beta}(k, n-k+1)$$

Beta分布 $B(\alpha, \beta)$ 的期望和方差公式是:

- $E(Z) = \frac{\alpha}{\alpha+\beta}$
- $D(Z) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

将 $\alpha = k$ 和 $\beta = n-k+1$ 代入，我们得到 $U(0,1)$ 样本的次序统计量的期望和方差:

- **期望**: $E(X_{(k)}) = \frac{k}{k+(n-k+1)} = \frac{k}{n+1}$
- **方差**: $D(X_{(k)}) = \frac{k(n-k+1)}{(n+1)^2(n+2)}$

这就是你原题中计算 $E(X_{(1)})$ 和 $E(X_{(n)})$ 的理论基础！

4. 从 $U(0,1)$ 到任意分布的推广（变换法）

我们可以利用 $U(0,1)$ 的结果来计算其他分布的次序统计量的性质。

- **概率积分变换**: 如果一个随机变量 Y 的CDF是 $F_Y(y)$，那么随机变量 $U = F_Y(Y)$ 服从 $U(0,1)$ 分布。
- **逆变换**: 如果 $U \sim U(0,1)$，那么 $Y = F_Y^{-1}(U)$ 就服从CDF为 F_Y 的分布。

这个关系对于次序统计量也成立：
如果 $Y_{(k)}$ 是来自 Y 的次序统计量， $U_{(k)}$ 是来自 $U(0,1)$ 的次序统计量，则 $Y_{(k)} = F_Y^{-1}(U_{(k)})$。

在你原题中，$X \sim U(\theta - 1/2, \theta + 1/2)$。
这是一个线性变换。令 $U \sim U(0,1)$，则 $X = (\theta + 1/2 - (\theta - 1/2))U + (\theta - 1/2) = U + \theta - 1/2$。
因此，$X_{(k)} = U_{(k)} + \theta - 1/2$。
所以：
$E(X_{(k)}) = E(U_{(k)} + \theta - 1/2) = E(U_{(k)}) + \theta - 1/2 = \frac{k}{n+1} + \theta - 1/2$
这就解释了我们是如何得到 $E(X_{(1)})$ 和 $E(X_{(n)})$ 的。

5. 联合分布与协方差

两个次序统计量 $X_{(i)}$ 和 $X_{(j)}$ $(i < j)$ 的联合PDF也可用类似的直观方法推导：
$f_{X_{(i)},X_{(j)}}(x,y) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(x)]^{i-1}[F(y) - F(x)]^{j-i-1}[1 - F(y)]^{n-j}f(x)f(y)$
对于 $x < y$。

对于 $U(0,1)$ 分布，这个公式可以用来计算协方差，其结果为：
$Cov(X_{(i)}, X_{(j)}) = \frac{i(n-j+1)}{(n+1)^2(n+2)}$ (for $i \leq j$)

这个协方差公式在计算 $D(\hat{\theta}_2) = D(\frac{1}{2}(X_{(1)} + X_{(n)}))$ 时被用到，因为：
$D(\frac{1}{2}(X_{(1)} + X_{(n)})) = \frac{1}{4}[D(X_{(1)}) + D(X_{(n)}) + 2Cov(X_{(1)}, X_{(n)})]$

总结：你需要记住的核心结论

对于一个来自 $U(0,1)$ 分布的样本的次序统计量 $X_{(k)}$：

1. **分布**: $X_{(k)} \sim \text{Beta}(k, n-k+1)$
2. **期望**: $E(X_{(k)}) = \frac{k}{n+1}$
3. **方差**: $D(X_{(k)}) = \frac{k(n-k+1)}{(n+1)^2(n+2)}$
4. **协方差**: $Cov(X_{(i)}, X_{(j)}) = \frac{i(n-j+1)}{(n+1)^2(n+2)}$ (for $i \leq j$)

对于其他分布，特别是线性变换（如均匀分布 $U(a,b)$ 或指数分布），可以通过变换法利用上述 $U(0,1)$ 的结论来求解。

希望这份详细的梳理能帮助你彻底理解次序统计量，并能熟练应用它们解决问题！

记忆技巧

当然！对于这种看起来复杂又没有明显规律的公式，最好的记忆方法不是死记硬背，而是**理解其背后的直观思想和模式**。一旦你抓住了核心的直观模型，这些公式就会变得非常自然，甚至可以当场推导出来。

下面我为你介绍一个非常有效的记忆框架，它基于一个简单的比喻："**分蛋糕**"或"**插旗子**"。

核心直观模型：在(0, 1)区间上插旗子

想象一下，区间 $(0,1)$ 是一块长度为 1 的长条蛋糕。你随机地往上面撒下 n 颗巧克力豆（这就是你的样本 $X_1,...,X_n$）。

$X_{(1)},X_{(2)},...,X_{(n)}$ 就是这些巧克力豆从左到右排好序的位置。

这 n 颗巧克力豆，会把整个蛋糕分成了 $n+1$ 个小段：

- 第1段：从 0 到 $X_{(1)}$
- 第2段：从 $X_{(1)}$ 到 $X_{(2)}$
- ...
- 第 k 段：从 $X_{(k-1)}$ 到 $X_{(k)}$
- ...
- 第 $n+1$ 段：从 $X_{(n)}$ 到 1

核心洞察：由于这些点是随机均匀撒下的，**平均来看，这 $n+1$ 段的长度是相等的**。
因为总长度是 1，所以每一小段的**平均长度**都是 $\frac{1}{n+1}$。

有了这个"平均每段长 $\frac{1}{n+1}$"的核心思想，我们来破解所有公式。

---

1. 期望 $E[X_{(k)}]$：第k个旗子的位置

公式：$E[X_{(k)}]=\frac{k}{n+1}$

记忆方法："第k个旗子的位置 = 前k段蛋糕的长度之和"

$X_{(k)}$ 这个点的位置，就是从 0 开始，数过 k 小段蛋糕的终点。
既然平均每段长度是 $\frac{1}{n+1}$，那么 k 段的总长度自然就是 $k\times\frac{1}{n+1}$。

- **例子**：$E[X_{(1)}]$? 第一个旗子的位置。就是第一段的长度，所以是 $\frac{1}{n+1}$。
- **例子**：$E[X_{(n)}]$? 最后一个旗子的位置。就是前 n 段的长度，所以是 $\frac{n}{n+1}$。

这个最简单，也最重要。

---

2. 方差 $D(X_{(k)})$ 和协方差 $Cov(X_{(i)}, X_{(j)})$：波动的程度

这两个公式看起来最吓人，但它们共享一个结构，特别是分母。

公式:

- $D(X_{(k)}) = \frac{k(n-k+1)}{(n+1)^2(n+2)}$
- $Cov(X_{(i)}, X_{(j)}) = \frac{i(n-j+1)}{(n+1)^2(n+2)}$ (其中 $i \le j$)

记忆方法 (分两步走):

第一步：记住那个“大魔王”分母

分母都是 $(n+1)^2(n+2)$。
你可以这样记：从期望的分母 $n+1$ 出发，先把它平方，再乘上比它大1的数 $n+2$。
口诀：“（n+1）平方，再乘（n+2）”

第二步：记住分子（非常有规律！）

方差的分子 $k(n-k+1)$：
口诀：“左边的段数 × 右边的段数”

- $X_{(k)}$ 这个点，把它左边的蛋糕段数数出来，正好是 k 段。
- 把它右边的蛋糕段数数出来，是 $(n+1)-k = n-k+1$ 段。
- 分子就是这两者相乘。
- **直观理解**：一个点的位置波动大小，和它两边空间的“自由度”有关。当点在中间时（$k \approx n/2$），两边空间都很大，波动最大，这个乘积也最大。当点在两端时（$k=1$ 或 $k=n$），一边空间被压缩，波动就小。

协方差的分子 $i(n-j+1)$：
口诀：“左边点的左边段数 × 右边点的右边段数”

- 对于 $X_{(i)}$ 和 $X_{(j)}$ (假设 $i \le j$)
- 取最左边的点 $X_{(i)}$，看它左边有几段：i 段。
- 取最右边的点 $X_{(j)}$，看它右边有几段：$n-j+1$ 段。
- 分子就是这两者相乘。
- **你会发现，方差是协方差在 $i=j=k$ 时的特例！** 所以你只需要记住协方差的规律就行了。

---

3. 极差的期望 $E[R] = E[X_{(n)} - X_{(1)}]$

公式：$E[R] = \frac{n-1}{n+1}$

记忆方法1：用期望公式硬算
$E[R] = E[X_{(n)}] - E[X_{(1)}] = \frac{n}{n+1} - \frac{1}{n+1} = \frac{n-1}{n+1}$。
这个最直接。

记忆方法2：继续用"分蛋糕"模型
口诀："内部蛋糕段的总长度"

- 极差 $R = X_{(n)} - X_{(1)}$ 是最外面那个点和最里面那个点之间的距离。
- 这段距离包含了多少个"小段"？它包含了从 $X_{(1)}$ 到 $X_{(2)}$，...，一直到 $X_{(n-1)}$ 到 $X_{(n)}$ 的所有段。总共有 $n-1$ 段。
- 既然平均每段长度是 $\frac{1}{n+1}$，那么 $n-1$ 段的总长度自然就是 $\frac{n-1}{n+1}$。

总结与表格

统计量	公式	记忆口诀/方法
期望 $E[X_{(k)}]$	$\frac{k}{n+1}$	**第k个位置 = k段蛋糕**
极差期望 $E[R]$	$\frac{n-1}{n+1}$	**内部n-1段蛋糕**
方差/协方差分母	$(n+1)^2(n+2)$	**(n+1)平方，再乘(n+2)**
方差分子 $D(X_{(k)})$	$k(n-k+1)$	**左边段数 × 右边段数**
协方差分子 $Cov(X_{(i)}, X_{(j)})$	$i(n-j+1)$	**左点的左边段数 × 右点的右边段数**

希望这个框架能帮你把这些看似杂乱的公式串成一个有逻辑、有图像的整体，从而轻松记住它们！