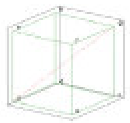


```
PlutoUI.Resource("https://tva1.sinaimg.cn/thumbnail/e6c9d24egy1h2alsw1tzxj20m80gomxn.jpg")
```

ch08 sec8.8 概率,平均数和中位数

```
md"""
# ch08 sec8.8 概率,平均数和中位数
"""
```

📕 Table of Contents

概率

假设我们随机从人群中挑选一个人, 问这个人出现在某个区间的可能性有多大, 比如说$70-80$岁之间, 根据$sec8.7$的数据, $70-80$岁之间占比在0.05. 我们可以定义:

$$从人群中随机挑选一个人，其年龄在\ 70-80\ 之间的概率为\ 0.05$$

这可以做一下理解: 如果一个年龄段的人占比越多, 那么从这个年龄段中随机挑选出的机会越大

积分定义为:

$$一个人\ 属于\ [a,b]\ 区间的概率为\ \ = \int_{a}^{b}p(t)dt$$

累积概率和上面的区间内定义一样,只不过区间初始点为0:

$$\ = \int_{0}^{b}p(t)dt$$

累积积分可以告诉我们比一个年龄年轻的人占总人群的比例

```
md"""
## 概率

假设我们随机从人群中挑选一个人, 问这个人出现在某个区间的可能性有多大, 比如说$70-80$岁之间, 根据$sec8.7$的数据, $70-80$岁之间占比在$0.05$. 我们可以定义:

$从人群中随机挑选一个人，其年龄在 70-80 之间的概率为 0.05$

这可以做一下理解: 如果一个年龄段的人占比越多, 那么从这个年龄段中随机挑选出的机会越大

积分定义为:

$一个人 属于 [a,b] 区间的概率为 \ = \int_{a}^{b}p(t)dt$

累积概率和上面的区间内定义一样,只不过区间初始点为$0$:

$\ = \int_{0}^{b}p(t)dt$

累积积分可以告诉我们比一个年龄年轻的人占总人群的比例




"""
```

Example

example 1

分析一下沿海小镇的捕鱼业情况,渔船每天最高捕捞8吨鱼,最少捕捞2d 吨鱼

- 利用概率密度函数如图的日常捕捞状况.解释其意义
- 捕捞$5-7$ 吨的概率有多大?

假设统计的是一年里的捕捞数据

累积函数分析了少于某个捕捞量的天数所占比例. 因为最少捕捞量为 2 吨. 所以累积函数为:

$$\int_2^t p(x)dx$$

与人口统计问题一样 $p(x)$ 描述了某个捕捞量所占总捕捞量的比例.

整个$p(x)$ 函数是分段函数在$2-6$ 区间是递增函数, 在 $5-8$ 之间是递减函数.

利用图中三个点的坐标, 可以计算出 两条直线的方程:

$$p(x)=\begin{cases} 0.04x\ , x\in[2,6] \\ -0.06x+0.6\ , x\in(6,8] \end{cases}$$

当 $t\in[2,6]$ 时,积分表示为:

$$\int_2^t 0.04xdx=0.02t^2-0.08$$

当 $t\in(6,8]$,由于两部分导数不同,需要用加法处理:

$$P(t)=\int_2^t p(x)dx=\int_2^6 p(x)dx+\int_6^t p(x)dx$$

带入化简得到:

$$\int_2^t p(x)dx=-0.03t^2+0.06t-1.88$$

因此总的累积函数表示为:

$$P(t)=\begin{cases} 0.02t^2-0.08\ , t\in[2,6] \\ -0.03t^2+0.6t-1.88\ , t\in(6,8] \end{cases}$$

累积分布的图如下,解释为, 捕捞量最大为 8吨. 小于8吨 就包括了所有能打到鱼的时间,所以为 1.

2. 5-7吨捕鱼量占总比可以表示为两个累积分布的差

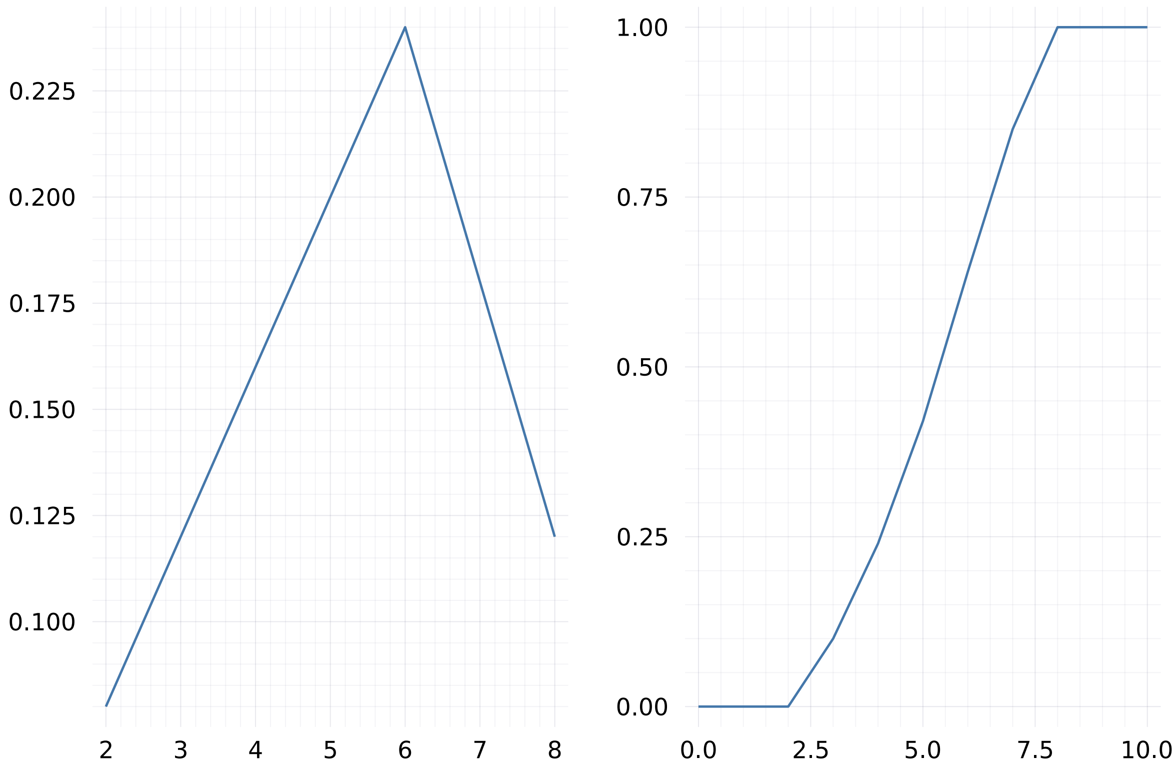
$$P(7) - P(5)$$

计算结果为 0.43,也就是说如果每年出海 100 天, 有 43 天的捕鱼量在$5-7$吨之间

可以看到微积分和概率理论有密切的联系.

```
md"""

!!! example
    example 1

    分析一下沿海小镇的捕鱼业情况,渔船每天最高捕捞$8$吨鱼,最少捕捞$2$d 吨鱼

    - 利用概率密度函数如图的日常捕捞状况.解释其意义
    - 捕捞$5-7$ 吨的概率有多大?


假设统计的是一年里的捕捞数据

累积函数分析了少于某个捕捞量的天数所占比例. 因为最少捕捞量为 2 吨. 所以累积函数为:

$\int_{2}^{t} p(x)dx$

与人口统计问题一样 $p(x)$ 描述了某个捕捞量所占总捕捞量的比例.

整个$p(x)$ 函数是分段函数在$2-6$ 区间是递增函数, 在 $5-8$ 之间是递减函数.

利用图中三个点的坐标, 可以计算出 两条直线的方程:

$p(x)=\left\{\begin{matrix}
 0.04x \ ,  x\in [2,6]\\
-0.06x+0.6 \ , x \in (6,8]
\end{matrix}\right.$

当 $t \in [2,6]$ 时,积分表示为:

$\int_{2}^{t}0.04x dx=0.02t^2-0.08$

当 $t \in (6,8]$,由于两部分导数不同,需要用加法处理:

$P(t)=\int_{2}^{t}p(x)dx=\int_{2}^{6}p(x)dx+\int_{6}^{t}p(x)dx$

带入化简得到:

$\int_{2}^{t}p(x)dx=-0.03t^2+0.06t-1.88$

因此总的累积函数表示为:

$$P(t)=\left\{\begin{matrix}
 0.02t^2-0.08\ , t\in [2,6]\\
-0.03t^2+0.6t-1.88 \ , t \in (6,8]
\end{matrix}\right.$$

累积分布的图如下,解释为, 捕捞量最大为 $8$吨.  小于$8$吨 就包括了所有能打到鱼的时间,所以为 1.


2. 5-7吨捕鱼量占总比可以表示为两个累积分布的差
```

```
$P(7)-P(5)$

计算结果为 0.43,也就是说如果每年出海 100 天, 有 43 天的捕鱼量在$5-7$ 吨之间


可以看到微积分和概率理论有密切的联系.


"""
```



```julia
let
  fspan=0:1:10
  function culumation(t)
      if  t<=2
         return 0
      elseif 2<t<=6
        return  0.02*(t^2)-0.08
      elseif   6<t<=8
        return  -0.03*(t^2)+(0.6*t)-1.88
      else
      return   1
      end
  end
  @show p5to7=culumation(7)-culumation(5)
  p1 = plot([2,6,8],[0.08,0.24,0.12],label=false,xticks=1:8,size=(600,400))
  p2=plot(culumation, fspan, label=false,size=(600,400))
  plot!(p1,p2)


end
```

```
p5to7 = culumation(7) - culumation(5) = 0.43000000000000055
```

中位数和平均数

中位数

在一个分布中,如果一个未知量x的取值满足一半种群数量的取值大于它, 另一半种群的取值小于它, 这时的取值T 就定义为中位数. 也就是累积积分值为 0.5

$$\int_{-\infty}^{T}p(x)dx=0.5$$

$$p(x)$$

为概率密度函数. 从图形角度看, 中位数把概率分布图分为两个面积相等的区域

平均和正态分布

这部分我们在统计学里专门讲

```
md"""
## 中位数和平均数


### 中位数

在一个分布中,如果一个未知量$x$的取值满足一半种群数量的取值大于它, 另一半种群的取值小于它, 这时的取值
$T$ 就定义为中位数.
也就是累积积分值为 0.5

$\int_{-\infty}^{T}p(x)dx=0.5$

$p(x)$ 为概率密度函数. 从图形角度看, 中位数把概率分布图分为两个面积相等的区域


## 平均和正态分布


这部分我们在统计学里专门讲

"""
```

```
let
@htl("""

<script>hljs.highlightAll();</script>
<script src="https://cdn.bootcdn.net/ajax/libs/mathjax/3.2.0/es5/tex-svg-
full.min.js">
<script src="http://127.0.0.1:8080/tex-svg-full.min.js"></script>
""")
end
```