

## گزارش پروژه‌ی شبیه سازی ۱ احتمال

### آریا ادیبی

قبل از هر چیز ذکر چند نکته ضروری است:

- اسمی که خواسته شده بود برای فایل‌های کد خود بگذاریم اسمی استاندارد نبود و من به دلیل این که مشکلی پیش نیاید یا حداقل «اخطار» این موضوع را نبینم این اسم را به `qi_studentNumber.m` تغییر دادم که در آن `i` شماره‌ی پرسش است.
- کامپیوتر من قدرت نسبتاً زیادی دارد از این رو نمی‌دانستم که شمار آزمایش‌ها یا متغیرهای مربوط دیگر را چقدر بگذارم که از کامپیوتر مقصد زمان زیادی نگیرد. از این رو در هر کد قسمتی به نام `initialization` قرار گرفته که بتوان تمامی متغیرهای لازم پرسش را به مقدار دلخواه عوض کرد. این انعطاف بیشتری هم به کد می‌دهد. برخی از پرسش‌ها هم شمار آزمایش‌ها را بیشتر از شمار گفته شده گذاشتم که نتیجه‌ی بهتری به دست آید. اگر لازم دیدید در همین قسمت به شمار دلخواه تغییر دهید.
- متغیرهای کدها به گونه‌ای اسم گذاری و «توضیح» گذاری شده‌اند که کد به راحتی قابل فهم باشد از این رو آن‌ها را زیاد توضیح نمی‌دهم.

### پرسش ۱ - تطابق تولدها!

کد روشن است و به خوبی «توضیح» گذاری شده است. تنها نکته‌ای که است این است که برای پیدا کردن کمینه‌ی  $n$  به طوری که احتمال مورد نظر بیشتر مساوی  $0.9$  شود از الگوریتم جست‌وجوی دودویی استفاده شده است. همچنین در این کد زمان حساب هر بخش هم چاپ می‌شود که اگر خواستید حذف کنید. جواب‌های خواسته شده را با اجرا کردن کد ببینید. نمودارهای حاصل را می‌توانید در شکل ۱ ببینید.

### پرسش ۲ - درهای جادویی

این پرسش نیز نام گذاری متغیرها و «توضیح» گذاری‌ها کد را روشن می‌کند. برای دیدن جواب کد را اجرا کنید. نتیجه‌ای که می‌گیریم این است که استدلال ما برای پرسش مشهور `TV game show` درست بود. یعنی احتمال بودن جایزه پشت آن پرده‌ای که حذف نشده و توسط ما انتخاب نشده برابر  $2/3$  و این احتمال برای پرده‌ی انتخاب شده  $1/3$  است. از این رو عاقلانس که پرده‌ی خود را عوض کنیم.

### پرسش ۳ - علی بی‌کار و سکه‌های پرتاب شونده

این پرسش نیز نام گذاری متغیرها و «توضیح» گذاری‌ها کد را روشن می‌کند. برای دیدن جواب کد را اجرا کنید.

برای حدس نظری این عدد آزمایش پرتاب سکه را در نظر بگیرید تا خط بیاید («موفقیت»). متغیر تصادفی  $X$  هندسی را به طور معمول تعریف کنید. داریم

$$\begin{aligned} P(X \geq r) &= \sum_{i=r}^{\infty} (1-p)^i - 1 \times p \\ &= \sum_{i=r}^{\infty} (1/2)^i = \left(\frac{1}{2}\right)^{r-1} \end{aligned}$$

با این که در آزمایش اصلی این کمینه  $r$  تا شیر پشت سر هم می‌تواند جاهای مختلفی بیاید ولی چون ۵۰ عدد بزرگی نیست این آزمایش تقریب خیلی بدی نخواهد بود. برای  $r = 6$  خواهیم که احتمال کمینه  $r$  تا شیر پشت هم بیاید  $\frac{1}{32}$  است که عدد کمی است. از این رو حدس می‌زنم که شمار مورد انتظار ما (E) کمتر مساوی ۶ باشد.

### پرسش ۴ - امید دگرگون

این پرسش نیز نام گذاری متغیرها و «توضیح» گزاری‌ها کد را روشن می‌کند. برای دیدن جواب کد را اجرا کنید. دلیل درستی عبارت به صورت زیر است:

$$\begin{aligned} P(X \geq r) &= \sum_{i=r}^{\infty} P(X = i) \\ \Rightarrow \\ \sum_{i=1}^{\infty} P(X \geq i) &= \sum_{i=1}^{\infty} i P(X = i) = \sum_{x \in A} x P(X = x) = E(X) \end{aligned}$$

### پرسش ۵ - غارهای شرطی

• در این پرسش فرض کردم اگر در درست انتخاب شود بلافاصله بیرون می‌رویم.

این پرسش نیز نام گذاری متغیرها و «توضیح» گزاری‌ها کد را روشن می‌کند. برای دیدن جواب کد را اجرا کنید.

### پرسش ۶ - توضیح مشترک دوگانه

رابطه‌ی گفته شده Law of Unconscious Statistician است. اگر شمار متغیرهای تصادفی ۱ باشد اثبات آن در کتاب آقای قهرمانی آورده شده است. من اثبات را برای ۲ متغیر تصادفی می‌آورم. اثبات برای شمار بیشتری متغیر تصادفی کاملاً مشابه است و من این کار را کردم که اثبات بدون دلیل شلوغ نشود.

تعریف. تابع  $f_{XY}$  را تابع احتمال توأم متغیرهای تصادفی  $X$  و  $Y$  بگیرید.

تعریف. متغیر تصادفی  $Z$  را برابر  $Z = h(X, Y)$  بگیرید.

می‌دانیم برای تمامی  $z \in \text{Domain}(Z)$  خواهیم داشت:

$$\begin{aligned} f_Z(z) &= P(Z = z) \\ &= \sum_{\{(x,y)|h(x,y)=z\}} f_{XY}(x,y) \end{aligned}$$

**تعریف.** به ازای هر  $z \in \text{Domain}(Z)$  بگیرید:  $A(z) = \{(x,y)|h(x,y) = z\}$

با تعریف‌ها و نکته‌ی گفته شده خواهیم داشت:

$$\begin{aligned} E[Z] &= \sum_z z f_Z(z) \\ &= \sum_z \sum_{A(z)} h(x,y) f_{XY}(x,y) \end{aligned}$$

این دو جمع معادل جمع روی تمام  $(x,y)$ ‌های ممکن است. پس:

$$E[Z] = \sum_{(x,y)} h(x,y) f_{XY}(x,y)$$

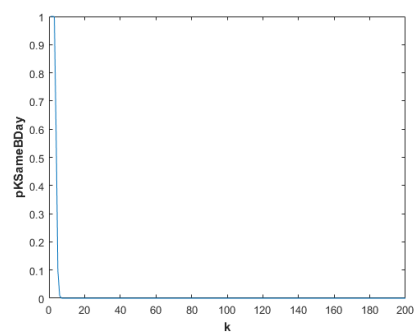
دقیقاً همان چیزی که می‌خواستیم. ■

این پرسش نیز نام گذاری متغیرها و «توضیح» گزاری‌ها کد را روشن می‌کند. برای دیدن جواب کد را اجرا کنید. خواهید دید که این ۲ مقدار را یکی تشخیص می‌دهد. نمودارهای حاصل را می‌توانید در [شکل ۲](#) ببینید.

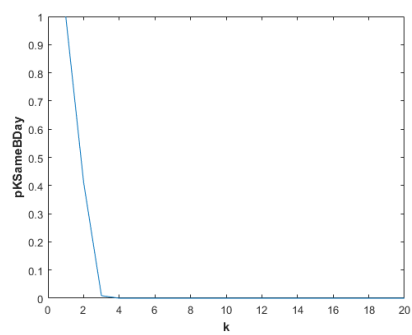
## پرسش ۷ - مجموع

طبق قضیه‌ی حد مرکزی حدس این است که با بیشتر شدن  $n$  رفتاری مشابه‌تر متغیر تصادفی نرمال داشته باشیم. که همانگونه که می‌بینیم همینگونه است.

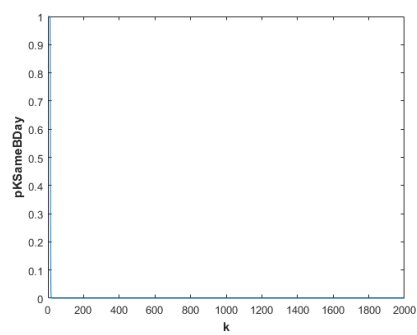
این پرسش نیز نام گذاری متغیرها و «توضیح» گزاری‌ها کد را روشن می‌کند. برای دیدن جواب کد را اجرا کنید. برخی از نمودارهای حاصل را می‌توانید در [شکل ۳](#) ببینید.



pKSameBDay vs K – n= 200 (ب)

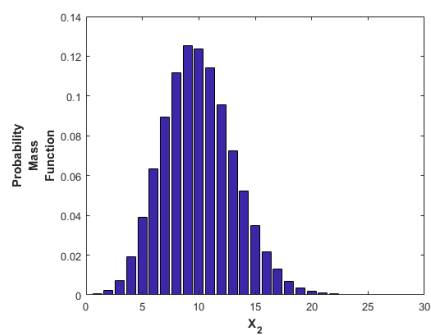


pKSameBDay vs K – n= 20 (ا)

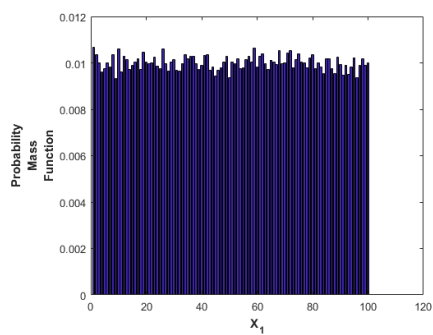


pKSameBDay vs K – n= 2000 (ج)

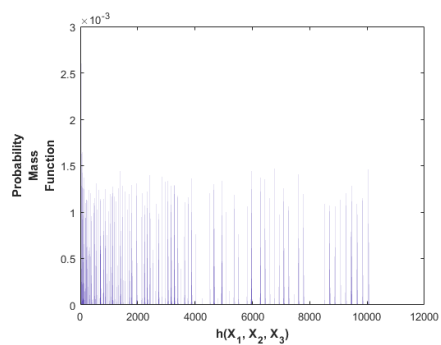
شکل ۱: نمودارهای پرسش ۱



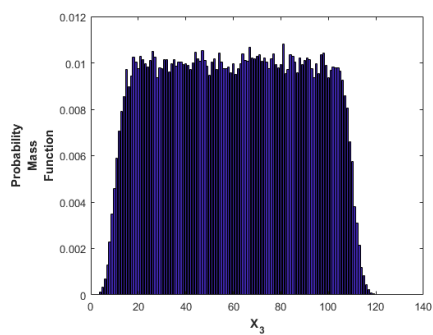
Probability mass function  $X_2$  (ب)



Probability mass function  $X_1$  (ا)

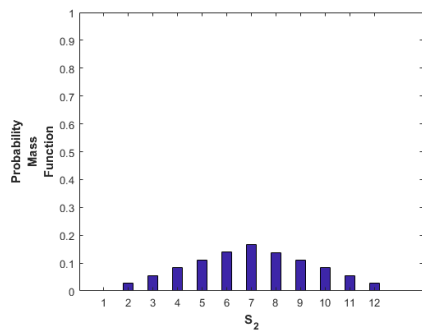


Probability mass function  $h(X_1, X_2, X_3)$  (د)

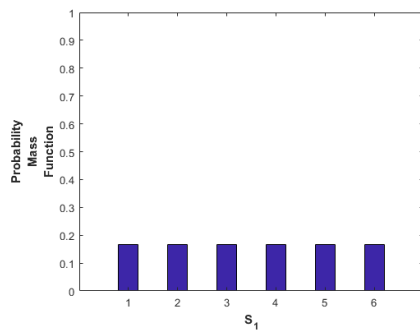


Probability mass function  $X_3$  (ج)

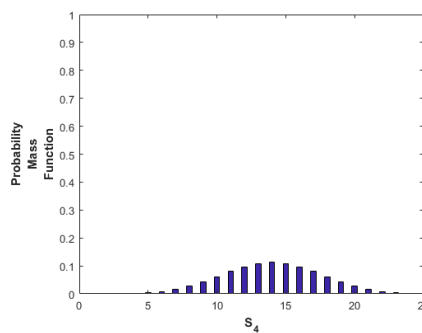
شکل ۲: نمودارهای پرسش ۶



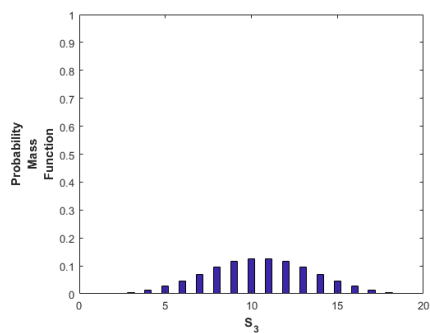
$n = 2$  (ب)



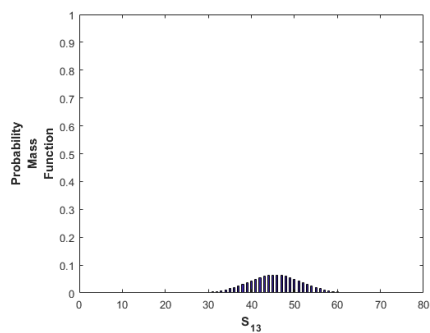
$n = 1$  (ا)



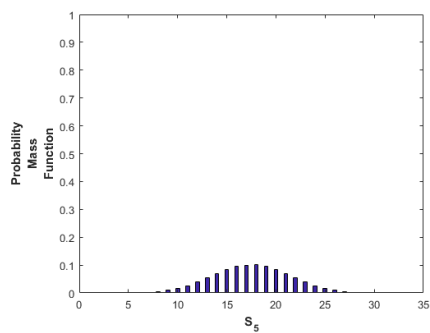
$n = 4$  (د)



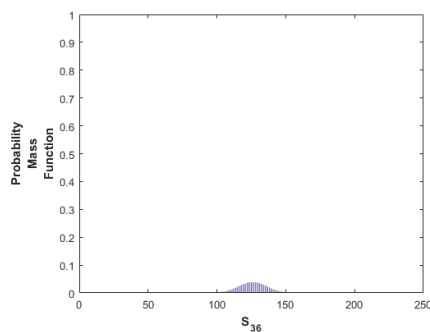
$n = 3$  (ج)



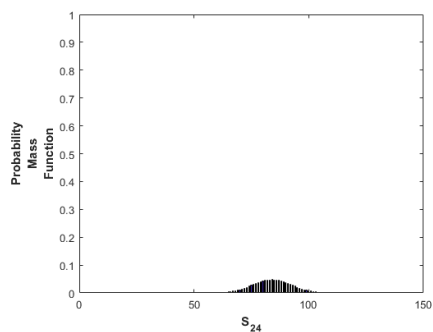
$n = 13$  (ه)



$n = 5$  (و)



$n = 36$  (ز)



$n = 24$  (ح)

۶