



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



گزارش تمرین شماره یک درس یادگیری تعاملی پاییز ۱۴۰۱

نام و نام خانوادگی
سیاوش رزمی
شماره دانشجویی
۸۱۰۱۰۰۳۵۲

فهرست

۳	چکیده
۴	سوال ۱ -
۴	هدف سوال
۵	سوال ۲ -
۵	هدف سوال
۵	توضیح پیاده سازی
۶	نتایج
۶	روند اجرای کد پیاده سازی

سوال ۳ -	۷
هدف سؤال	۷
توضیح پیاده‌سازی	۷
سوال ۴ -	۱۱
هدف سؤال	۱۱
توضیح پیاده‌سازی	۱۱
نتایج	۱۱
سوال ۵ -	۱۲
هدف سؤال	۱۲
سوال ۶ -	۱۳
هدف سؤال	۱۳
نکات مهم و موارد تحویلی	۱۴
موارد تحویلی	۱۴
منابع	۱۵

چکیده

هدف کلی این تمرین آشنایی با مفاهیم آماری و ابزارهای مبتدی پایتون است.

برای بررسی میزان اثر بخشی داروی کاهش فشار خون با دو فرمولاسیون مختلف از سه استراتژی مختلف تجویز بر روی بیمارها استفاده شده است که با بررسی نتایج قصد داریم میزان اثر بخشی دو فرمولاسیون را بررسی کنیم. برای بررسی این استراتژی ها ابتدا میانگین پاداش هر کدام را در هر trial رسم کرده و به شکل شهودی آنها را بررسی میکنیم، سپس با استفاده از نمودار box-plot میزان پاداش آخرین بیمار را در ۱۰ مرتبه تجویز مقایسه میکنیم و سپس با استفاده از آزمون های آماری فرض متفاوت بودن اثر بخشی دو فرمولاسیون را با یکدیگر را آزمایش میکنیم، در آخر نیز با بررسی این استراتژی ها با روش A/B testing مزایا و معایب هر استراتژی را بیان میکنیم.

سوال ۱ -

هدف سوال

هدف این سؤال بررسی محیط آزمایش و نحوه گرفتن پاداش است.

جواب سوال:

- به دلیل اینکه هر بدن هر شخص ممکن است به فرمولاسیون های مختلف واکنش های متفاوتی نشان بدهد و میزان تأثیر دارو بر هر فرد متفاوت است، همچنین در طول فرآیند آزمایش نیز ممکن است خطای انسانی و اندازه گیری در فرایند تولید دارو، تجویز یا اندازه گیری میزان پاداش دخیل باشد که این موضوع بر روی نتایج دریافت شده تأثیر می گذارد.

سوال ۲ -

هدف سوال

هدف سؤال این است که با پیاده‌سازی هر ۳ استراتژی برروی ۱۰۰ بیمار و رسم نمودار میانگین پاداش در هر مرتبه میزان پاداش در هر استراتژی را تحلیل و مقایسه کنیم، هدف این کار میتواند درک فرایند تصادفی و بررسی این موضوع است که در این فرآیند ها میزان پاداش خروجی ما در هر مرتبه متفاوت است، همچنین عمل کرد این ۳ استراتژی در میزان کاهش فشار خون در این مسأله بررسی می‌شود.

توضیح پیاده سازی

در ابتدا ۳ تابع با نام های `doctor_a`, `doctor_b` و `doctor_c` تعریف شده‌اند که با فراخوانی هر کدام طبق استراتژی های تصمیم‌گیری هر پزشک که در صورت سؤال تعریف شده است ۱۰۰ مرتبه اقدام به انجام اکشن و ذخیره پاداش میکند.

تابع `doctor_a`:

اکشن اول به شکل تصادفی انتخاب شده سپس برای یک حلقه ۱۰۰ تایی:

۱- مقدار پاداش با استفاده از تابع `get_reward` محاسبه شده

۲- شماره `trial`، اکشن انجام شده و میزان پاداش ذخیره شده

۳- اگر پاداش مثبت بود به احتمال ۰.۸ همان اکشن و به احتمال ۰.۲ اکشن دیگر انتخاب می‌گردد

۴- اگر پاداش منفی بود به احتمال ۰.۳ همان اکشن و به احتمال ۰.۷ اکشن مقابل انتخاب می‌گردد

۵- در صورتی که پاداش برابر با صفر باشد اکشن به شکل تصادفی انتخاب می‌شود

۶- حاصل جمع تمام پاداش ها تا هر `trial` محاسبه می‌گردد.

۷- تمام مقادیر در یک آرایه ذخیره شده و به شکل خروجی داده می‌شود

تابع `doctor_b`:

در یک حلقه ۱۰۰ تایی:

۱- اکشن به شکل تصادفی انتخاب می‌شود

۲- پاداش با استفاده از تابع `get_reward` محاسبه شده

۳- مقادیر شماره `trial` و اکشن انجام شده و میزان پاداش ذخیره می‌شود

۴- حاصل جمع تمام پاداش ها تا هر `trial` محاسبه می‌گردد.

۵- تمام مقادیر در یک آرایه ذخیره شده و به شکل خروجی داده می‌شود

تابع `doctor_c`:

۱- برای ۱۰ مرتبه اکشن ۱ انجام شده و مقادیر آن ذخیره می‌شود

۲- برای ۱۰ مرتبه اکشن ۲ انجام شده و مقادیر آن ذخیره می‌شود

۳- اکشنی که بیشترین میزان پاداش را داشته باشد مشخص می‌شود

۴- در یک حلقه while تا زمانی که تعداد trial ها به ۱۰۰ برسد:

۱- هفت مرتبه اکشنی که بیشترین میزان پاداش را داشته اجرا می‌شود و مقادیر آن ذخیره می‌گردد

۲- سه مرتبه اکشن به شکل تصادفی انتخاب می‌گردد و مقادیر آن ذخیره می‌شود

۳- اکشنی که بیشترین میزان پاداش را داشته انتخاب می‌گردد

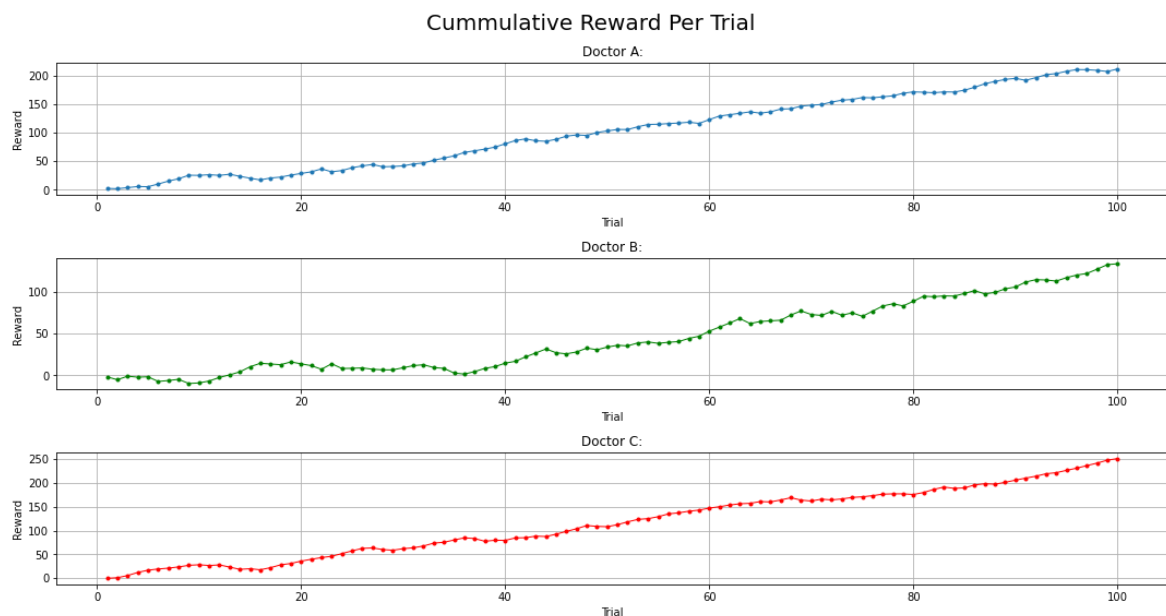
۵- حاصل جمع تمام پاداش ها تا هر trial محاسبه می‌گردد..

۶- تمام مقادیر در یک آرایه ذخیره شده و به شکل خروجی داده می‌شود

سپس در بخش اول هر کدام از توابع یکبار اجرا شده و نتایج آنها ذخیره می‌شود، سپس نمودارهای خواسته شده با استفاده از داده‌های خروجی رسم می‌شوند.

نتایج

نتایج به دست آمده از توابع شکل تصادفی دارند و در هر مرتبه اجرا ممکن است متفاوت نسبت به مرتبه قبل باشند، اما بطور کلی به نظر می‌رسد استراتژی اول و سوم حدوداً از استراتژی دوم بهتر عمل می‌کنند،



شکل ۱-۲: نمودار میانگین پاداش در هر trial با یک بار اجرا کد

روند اجرای کد پیاده‌سازی

توابع ذکر شده در یک فایل به نام strategy.py ذخیره شده‌اند و کدهای رسم نمودار در فایل Question*.ipynb ذخیره شده‌اند، بایستی فایل‌های strategy.py، BP_reward.py و فایل Question*.ipynb در یک Directory قرار داشته باشند، برای اجرای کد کافیست فایل Question1.ipynb را اجرا نمایید.

سوال ۳ -

هدف سؤال

در این سؤال با ماهیت فرآیند های تصادفی و نحوه بررسی و مقایسه ی این نوع فرآیند ها می پردازیم.

بعلت اینکه تابع پاداش محیط ما به شکل تصادفی است در هربار اجرا مقادیر متفاوتی تولید میکند که باعث می شود نمودار در هربار اجرا مقداری متفاوت باشد، اما با چندین مرتبه اجرای فرآیند، مشخص می شود که در عین تصادفی بودن این مقادیر از یک رفتار مشخصی پیروی میکنند که میتوان از آن برای تحلیل و مقایسه فرآیندها با یکدیگر استفاده کرد.

توضیح پیاده سازی

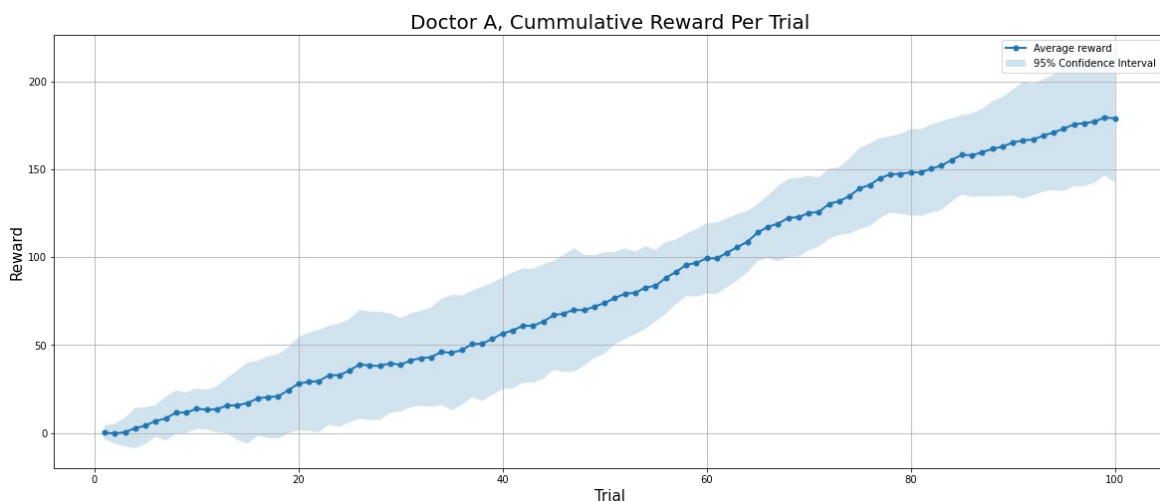
ابتدا به تعداد خواسته شده هر یک از استراتژی ها اجرا شده و نتایج آن ذخیره می شود، سپس با کنار هم گذاشتن مقادیر میانگین پاداش، میانگین و انحراف معیار هر trial حساب می شود، به دلیل کم بودن تعداد نمونه ها (۵ و ۲۰) از توزیع t برای محاسبه بازه اطمینان استفاده می کنیم با استفاده از فرمول زیر بازه اطمینان برای هر سه استراتژی حساب شده و نتایج رسم می شوند.

$$CI = \bar{X} \pm t^* \frac{s}{\sqrt{N}}$$

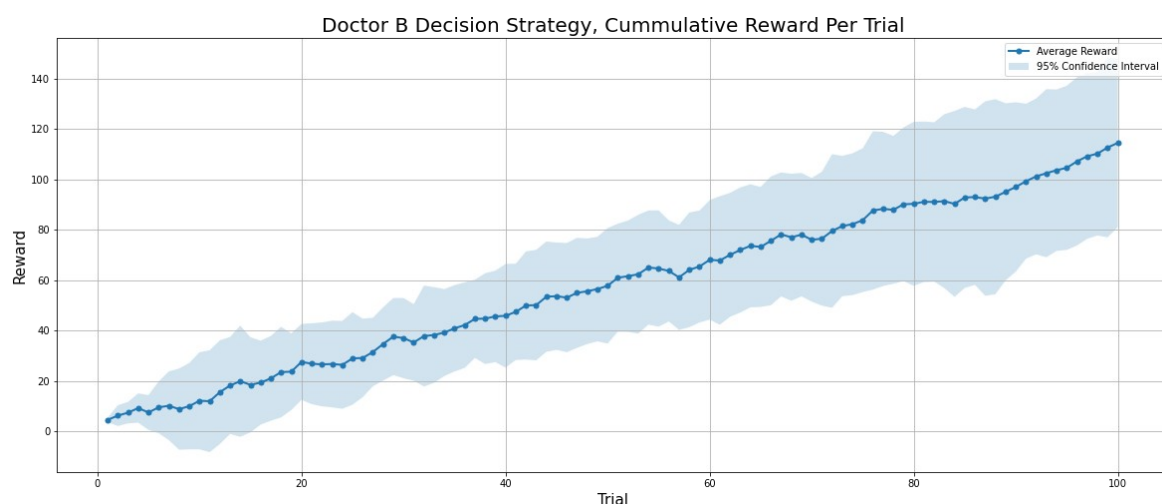
برای محاسبه t^* از کتابخانه ی `scipy.stats` استفاده شده است.

نتایج

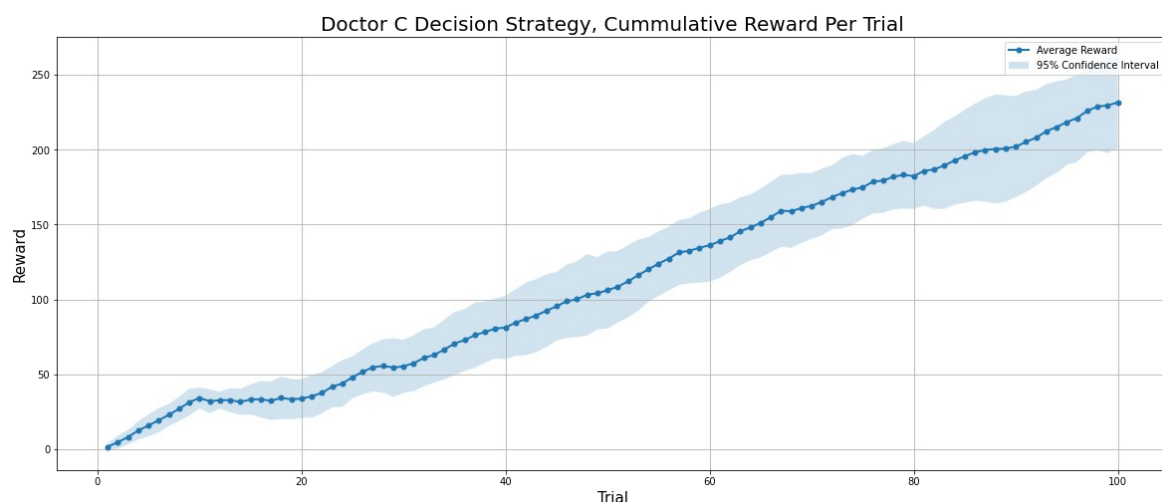
نتایج به دست آمده با ۵ مرتبه تکرار به شکل زیر است:



شکل ۳-۱: نمودار میانگین پاداش به trial استراتژی پزشک اول با ۵ مرتبه تکرار



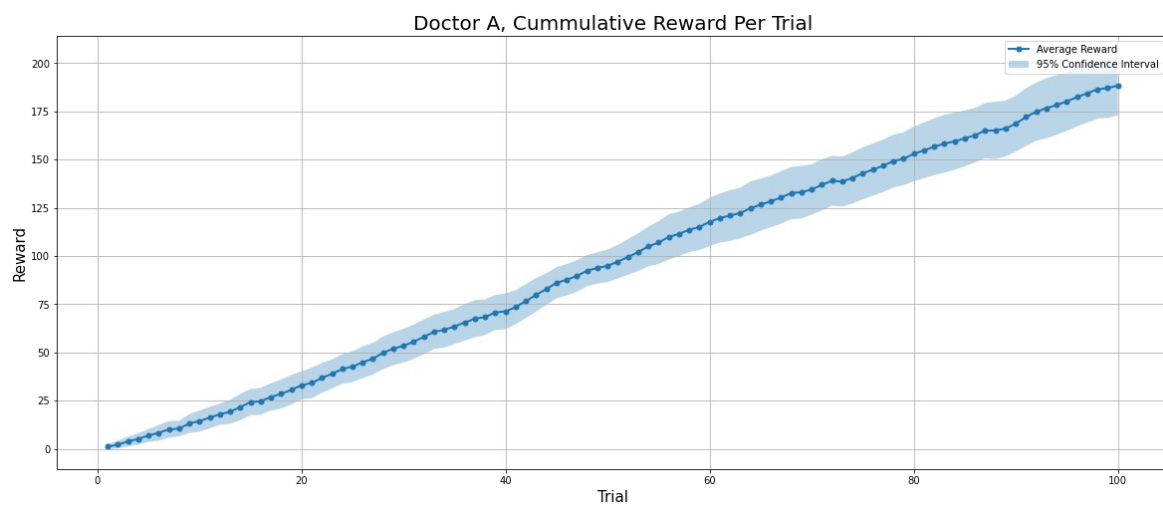
شکل ۲-۳: نمودار میانگین پاداش به trial استراتژی پزشک دوم با ۵ مرتبه تکرار



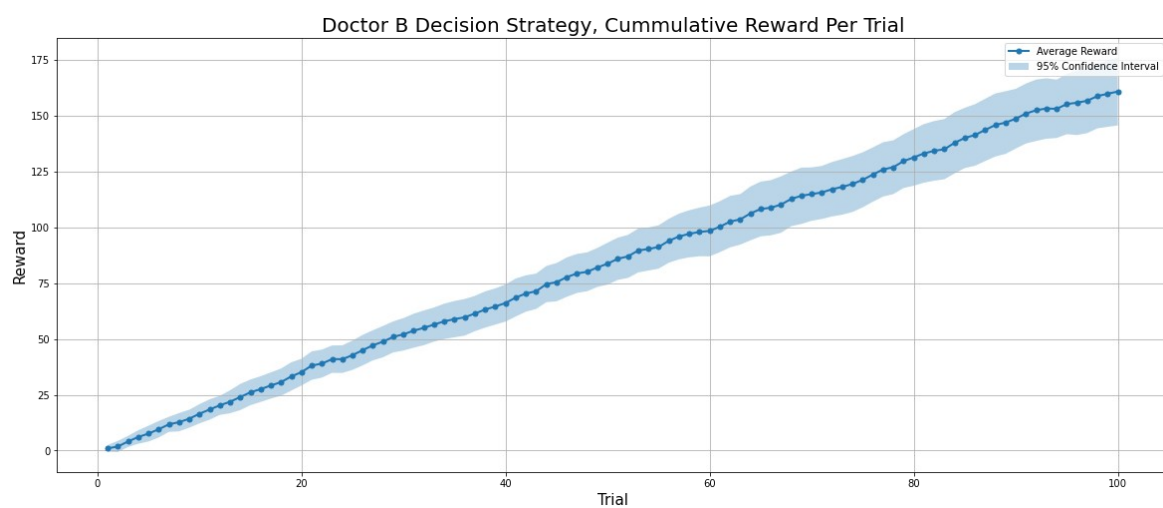
شکل ۳-۳: نمودار میانگین پاداش به trial استراتژی پزشک سوم با ۵ مرتبه تکرار

به نظر می‌رسد که با جمع کردن نتایج اجراها، رفتار نمودار به شکل پایدار تری رسیده است و نمودار در چندین مرتبه اجرای کد خروجی تقریباً یکسانی دارد، به نظر می‌رسد بازدهی استراتژی سوم از مابقی بهتر بوده و بیشترین حاصل جمع پاداش را در انتهای آزمایش پیدا می‌کند استراتژی اول نیز با مقدار کمی تفاوت از استراتژی سوم رتبه دوم را دارد، از لحاظ بازه اطمینان نیز تمام نمودارها در ابتدا بازه اطمینان محدودی دارند و هرچه جلوتر می‌رویم بازه اطمینان بزرگ‌تر می‌شود به این معنا که نتایج هر ۵ بار تکرار از حدوداً یک نقطه شروع می‌شود اما هر چه جلو می‌رویم نمودار نتایج هر تکرار مسیر متفاوتی پیدا می‌کند. نمودار استراتژی دوم بزرگترین بازه اطمینان را دارد که دلیل آن تصادفی بودن نحوه انتخاب اکشن است که باعث می‌شود در هر تکرار نتایج کاملاً متفاوتی اخذ شود.

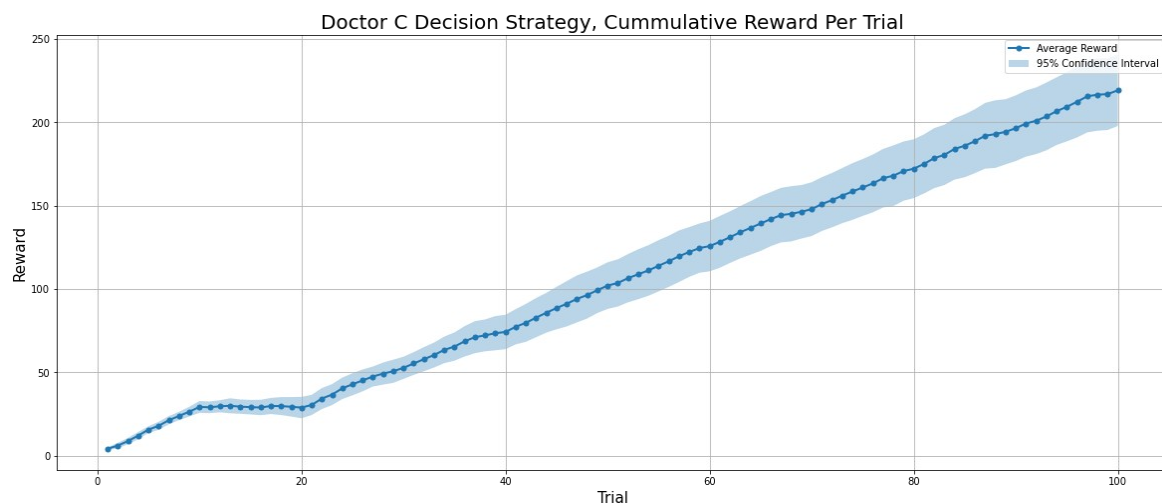
نتایج به دست آمده با ۲۰ مرتبه تکرار به شکل زیر است:



شکل ۴-۳: نمودار میانگین پاداش به trial استراتژی پزشک اول با ۲۰ مرتبه تکرار



شکل ۵-۳: نمودار میانگین پاداش به trial استراتژی پزشک دوم با ۲۰ مرتبه تکرار



شکل ۶-۳: نمودار میانگین پاداش به trial استراتژی پزشک سوم با ۲۰ مرتبه تکرار

با ۲۰ مرتبه تکرار فرآیند به نظر می‌رسد تفاوت میانگین پاداش استراتژی‌ها به شکل معنادارتری قابل مشاهده است، استراتژی سوم در این نمودار به شکل قابل توجهی از دو استراتژی دیگر بهتر است و استراتژی دوم که به شکل تصادفی انجام شد پایین‌ترین پاداش را دارد، همچنین بازه اطمینان در این نمودار‌ها باریک‌تر شده است که به این معناست که با انجام هر چه بیشتر آزمایش تخمین ما از میانگین پاداش استراتژی‌ها با قطعیت بیشتری قابل پیش‌بینی است.

سوال ۴ -

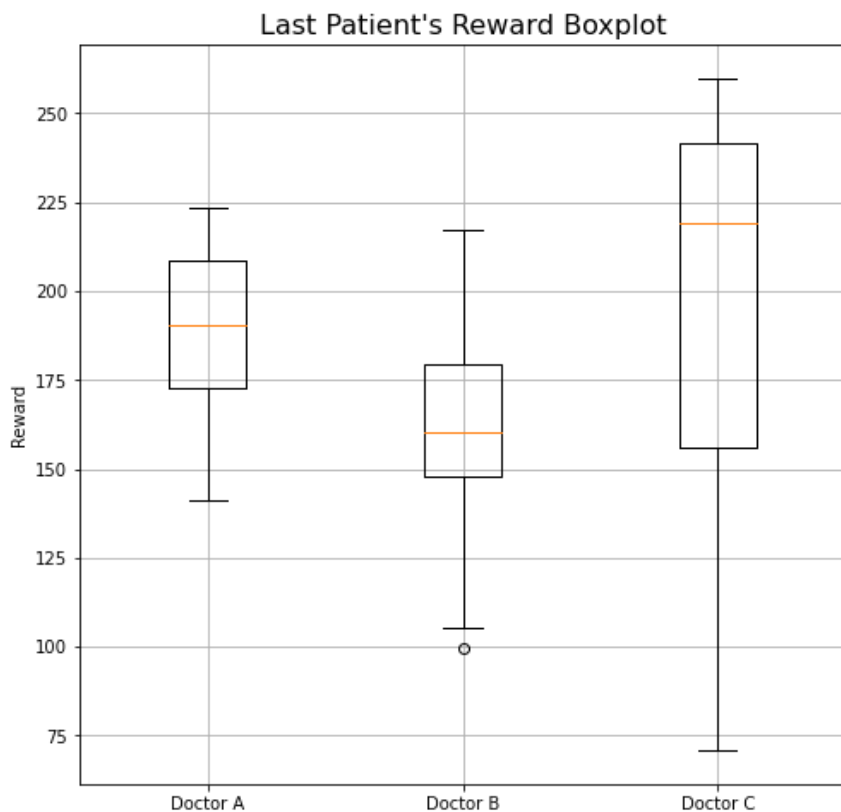
هدف سؤال

رسم نمودار box-plot برای بیمار آخر در ۱۰ بار تکرار، با رسم این نمودار تفاوت عمل کرد استراتژیها به شکل بهتری قابل مشاهده است.

توضیح پیاده‌سازی

پس از ۱۰ بار تکرار استراتژی ها مقدار آخرین trial آن‌ها در یک آرایه ذخیره شد و با استفاده کتابخانه ی matplotlib رسم شدند.

نتایج



شکل ۴-۱: نمودار box-plot نتایج میانگین پاداش آخرین بیمار در هر استراتژی با ۱۰ مرتبه تکرار

همانطور که از نمودار مشخص است استراتژی C به طور کلی از دو استراتژی دیگر و استراتژی A از حالت کاملاً تصادفی بهتر عمل کرده است، همچنین استراتژی C واریانس بیشتری داشته و مقادیر outlier در آن وجود دارد، البته نتایج با توجه به ماهیت تصادفی فرآیند ها در هر مرتبه اجرا مقداری متفاوت هستند ولی ترتیب میانه و چارک اول و سوم، بین نمودارها در تمام اجراها حدوداً مشابه است.

سوال ۵ -

هدف سؤال

هدف سؤال بررسی اثر بخشی دو دارو با استفاده از آزمون آماری است.

توضیح پیاده‌سازی

برای انجام این آزمون ابتدا هر اکشن ۱۰۰ مرتبه اجرا شد و مقادیر پاداش آنرا ذخیره کردیم، سپس میانگین و انحراف معیار مقادیر محاسبه شدند، برای انجام آزمون آماری از t-test استفاده کردیم (با توجه به مقادیر نمونه‌ها و درجه آزادی ۹۹ نتایج این تست با آزمون Z حدوداً برابر خواهد بود)

با توجه به شرایط آزمایش پیش فرض های آزمون t بررسی شدند:

۱- وابستگی درون و بین گروهی: تمام آزمایش ها به شکل مستقل و تصادفی بر روی افراد متفاوت (به شکل non-paired) انجام شده است بنابراین شرط برقرار است.

۲- تعداد نمونه: تعداد ۱۰۰ نفر برای هر اکشن در نظر گرفته شده است.

فرض صفر: میانگین پاداش گروه اول و گروه دوم برابر است.

فرض جایگزین: میانگین گروه اول از گروه دوم بیشتر است.

بر اساس فرض مورد آزمایش نتیجه میگیرم که تست یکطرفه است، برای آزمایش فرض صفر میتوان p-value را محاسبه کرد و یا از بازه اطمینان استفاده کرد، هر دو روش استفاده شد.

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0.37$$

$$T_{score} = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{SE} = 9.37$$

$$df = \min(n_1 - 1, n_2 - 1) = 99$$

کد مسأله در فایل Question5 قرار داده شده است.

سپس با استفاده از تابع t.sf مقدار p value محاسبه شد که برابر با $1.2478330233722415e^{-12}$ شد که با در نظر گرفتن مقدار آلفا برابر ۵ درصد، فرض صفر با قطعیت رد می شود.

بازه اطمینان نیز با استفاده از فرمول زیر محاسبه می‌شود:

$$(\bar{X}_1 - \bar{X}_2) \pm t^* SE = (2.34, 3.57)$$

با توجه به اینکه بازه اطمینان شامل مقادیر بزرگتر از صفر است و صفر را شامل نمی‌شود بنابراین میتوان فرض صفر را با این روش نیز رد کرد.

سوال ۶ -

هدف سؤال

بررسی و مقایسه روش A/B testing با سیاست های تقویتی

پاسخ:

در روش A/B ما به شکل ما به دو گروه مختلف با تعداد یکسان در شرایط کاملاً مشابه به شکل تصادفی دو اکشن متفاوت را اعمال میکنیم، با توجه به این توضیح میتوان گفت استراتژی پز شک دوم (کاملاً تصادفی) به روش A/B test شبیه تر است چون در این روش نیز ما به شکل تصادفی به حدود نصف بیماران یک دارو و به مابقی آن ها داروی دیگر را تجویز می کنیم.

مزایای این روش نسبت به سیاست های تقویتی این است که به دلیل اینکه ما دقیقاً یک اکشن مشخص را در دو گروه کاملاً یکسان آزمایش میکنیم به راحتی میتوان تشخیص داد کدام یک عمل کرد بهتری دارد، همچنین با این روش میتوان فرضیه های بسیار مشخص و خاص را آزمایش کرد، فرایند آزمایش نیز بسیار ساده تر از روش های تقویتی است.

معایب این روش اما این است که تنها یک فرضیه مشخص را که قابل اندازه گیری باشد با این روش قابل آزمایش است، فرایند آزمایش بسیار زمان بر و پرهزینه است و در صورتی که بخواهیم فرض دیگری را امتحان کنیم بایستی آزمایش را به طور کامل تکرار کنیم.

نکات مهم و موارد تحویلی

لازم است که به نکات زیر در نوشتن گزارش توجه داشته باشید.

۱. ساختار کلی گزارش که در این فایل به آن اشاره شده باید رعایت شود. در صورت تمایل می‌توانید از latex یا هر نرم افزار دلخواه دیگر برای نوشتن گزارش استفاده کنید، به شرط اینکه ساختار کلی گفته شده رعایت شود. لذا در صورت رعایت نکردن ساختار کلی گزارش بخشی از نمره تمرین کم خواهد شد.
۲. برای تصاویر و جداول موجود در گزارش حتما کپشن قرار داده شود.
۳. **نتایج و تحلیل‌های** شما در روند نمره دهی اهمیت بسیار بالایی دارد، لذا خواهشمندیم کلیه نتایج و تحلیل‌های خواسته شده به صورت کامل و دقیق در گزارش آورده شوند.
۴. در صورت مشاهده شباهت بین گزارش شما و افراد مختلف نمره این سری تمرین برای شما در نظر گرفته نمی‌شود.

موارد تحویلی

۱. برای هر سری از تمرینات، فقط یک فایل با فرمت PDF آماده کنید.
۲. به همراه فایل گزارش، یک پوشه به نام Codes ایجاد کنید و کدها و فایل‌های پیاده‌سازی هر سوال را به صورت تفکیک شده در پوشه‌های جداگانه قرار دهید.
۳. هیچ گونه جدول یا تصویر به صورت جداگانه خارج از گزارش ارسال نشود. مگر اینکه به صورت صریح در تمرین از شما خواسته شده باشد.
۴. در انتها، لطفاً برای هر تمرین گزارش و پوشه کدها را به صورت گفته شده، در یک فایل زیپ با فرمت زیر در سامانه یادگیری الکترونیک بارگذاری نمایید.

HW#_LastName_StudentNumber.zip

به طور مثال:

HW1_Mesbah_810111111.zip

منابع

در این بخش منابع (مقالات، سایت‌ها و ...) که در تمرینات و پیاده‌سازی استفاده کرده‌اید را به یک فرمت استاندارد برای مثال فرمت IEEE وارد کنید. برای رفرنس دهی می‌توانید از بخش Reference نرم‌افزار Word یا [این سایت](#) یا افزونه‌های دیگر استفاده کنید. لازم است منبع مورد استفاده خود را در بخش مربوطه ارجاع دهید.

توجه به این نکته ضروری است که میزان شباهت کد یا راه حل شما در صورت استفاده از منابع دیگر باید به حد معقولی باشد و کپی کردن از منابع مد نظر ما نیست. در صورت مشاهده کپی از منابع به صورت کامل نمره تمرین یا بخش مورد نظر به شما تعلق نمی‌گیرد.