University of Tehran

آمار و احتمالات مهندسی

تمرین هفتم – آمار

نرجس و دریا

سؤال ۱.

یک آژانس مسافرتی می‌داند که در طولانی مدت، ۹۰ درصد مسافرانی که صندلی رزرو می‌کنند در سفر خود حاضر می‌شوند. در یک سفر خاص با ۳۰۰ صندلی، آژانس مسافرتی ۳۲۴ رزرو را می‌پذیرد. برای پاسخ به سوالات زیر، از تقریب نرمال بهره بگیرید:

آ) آیا این احتمال وجود دارد که سفر بیش از حد مسافر داشته باشد؟ رویداد حضور هر مسافری مستقل از یکدیگر است.

ب) باتوجه به بخش قبلی، اکنون فرض کنید مسافران همیشه دو نفره سفر می‌کنند و تنها در صورتی سفر می‌کنند که هر دو نفربه صورت جفت حاضر شوند. بررسی کنید که آیا پاسخ شما به سوال قبلی با این نتیجه مطابقت دارد یا خیر.

سؤال ۲.

ما علاقه مند هستیم که درصد فارغ التحصیلان یک دانشگاه را که ظرف یک سال پس از اتمام دوره کارشناسی خود کار پیدا کرده‌اند، تخمین بزنیم. فرض کنید ما یک نظرسنجی انجام می‌دهیم و متوجه می‌شویم که ۳۴۸ نفر از ۴۰۰ فارغ التحصیل نمونه تصادفی شغل پیدا کرده‌اند. کلاس فارغ التحصیلی مورد بررسی شامل بیش از ۴۵۰۰ دانش آموز بود.

آ) پارامتر مطلوب تخمین مذکور را بدست آورید. ارزش تخمین نقطه‌ای این پارامتر چقدر است؟

ب) بررسی کنید که آیا شرایط ایجاد بازه‌ی اطمینان بر اساس این داده‌ها وجود دارد یا خیر.

ج) بازه اطمینان ۹۵ درصدی را برای نسبت فارغ التحصیلانی که در مدت یک سال پس از اتمام دوره کارشناسی خود در این دانشگاه پیدا کرده‌اند محاسبه کرده و آن را در چارچوب داده‌ها تفسیر کنید.

د) بدون محاسبه ی بازه اطمینان ۹۹ درصدی، عرض بازه‌ی محاسبه شده در قسمت ج را با عرض بازه ۹۹ درصدی مقایسه کنید. توضیح دهید کدام یک عریض‌تر است و چرا.

سؤال ۳.

یک شرکت بیمه دارای n مشتری بیمه شده است. فرض کنید که n عدد بسیار بزرگی است. کارشناسان بیمه با مطالعه داده‌های سالیان متمادی دریافته‌اند که رفتار بیمه شدگان در طول یک سال به صورت زیر قابل ساده سازی است: هر بیمه شده مستقل از سایرین با احتمال $1-p$ خرجی روی دست بیمه نمی‌گذارد. اما به احتمال p دچار حادثه شده و سقف پرداخت جبرانی بیمه که c تومان است را دریافت می‌کند. مقادیر p, c معلوم فرض شود.

واضح است که متوسط هزینه تحمیل شده به بیمه مقدار pc تومان به ازای هر فرد در سال است. لذا هیئت مدیره تصمیم گرفته است که برای کل سال جاری مقدار $(1+\eta)pc$ تومان را از هر بیمه شده تحت عنوان حق بیمه دریافت کند. به دلیل رقابت با سایر شرکت های بیمه این شرکت بطور خاص در امسال دنبال سود نیست اما ضرر نیز نمی‌خواهد.

حداقل بیمه شدگان تحت حمایت این شرکت، یعنی n، چقدر باشد تا به ازای مقدار مشخص η, p, c احتمال ضرر دهی شرکت کمتر از مقدار مفروض δ شود؟ (راهنمایی: اگر X توزیع نرمال استاندارد و t یک ثابت باشد: $P(X>t) \leq e^{\frac{-t^2}{2}}$)

سؤال ۴.

آزمایش‌هایی با هدف بررسی مقدار فلزی که در درجهٔ حرارت 120° حمام اسید از یک سری قطعات جدا می‌شوند انجام داده‌ایم. نتایج زیر به دست آمده اند. (هر عدد معرف ضخامت فلز جدا شده برحسب ۰٫۰۱ میلی متر است و داده‌ها را می‌توان دارای توزیع نرمال فرض کرد.)

$$2/4,\ 2/2,\ 2/0,\ 1/8,\ 2/1,\ 2/1$$

آ) بازه‌های اطمینان ۰٫۹۰ و ۰٫۹۵ برای میانگین مقدار فلزی که در شرایط فوق از قطعات جدا می‌شوند، به دست آورید.

ب) فرض کنید بجای داده‌ی ۲٫۲، مقدار ۳٫۱ مشاهده می‌شد. بازه‌های بند الف را دوباره به‌دست آورید، و درباره‌ی تاثیر این دادهٔ دور افتاده (یا داده‌ی پرت) در بازه‌های اطمینان بحث کنید.

سؤال ۵.

در یک نظرسنجی در مورد میزان رضایت دانشجویان یک دانشگاه از آموزش مجازی، ۶۳ درصد از آن‌ها اعلام کردند که از این نوع آموزش رضایت ندارند. حاشیه خطای این نظرسنجی برای بازه اطمینان ۹۵ درصد برابر با $\pm$۳٪ است.

آ) اندازه‌ی نمونه‌ی مورد استفاده در این نظرسنجی را تخمین بزنید.

ب) با استفاده از اندازه‌ی بدست آمده، یک بازه اطمینان ۹۰٪ برای درصد دانشجویانی که از آموزش مجازی ناراضی هستند بسازید.

سؤال ۶.

میانگین زمان پاسخگویی و انحراف معیار در یک سیستم کامپیوتری چندکاربره به ترتیب ۱۵ ثانیه و۳ ثانیه میباشد. احتمال آنکه زمان پاسخگویی بیش از ۵ ثانیه از میانگین فاصله داشته باشد را محاسبه کنید.

سؤال ۷.

یک منبع در هر زمانی که اندازه‌گیری می‌شود K فوتون از خود ساطع می‌کند. ما فرض می‌کنیم که K دارای توزیع زیر است:

$$p_K(k;\theta) = c(\theta)e^{-k\theta} \qquad k = ۰, ۱, ۲, ...$$

بطوری که θ معکوس دمای منبع است و $c(\theta)$ ضریب نرمالیزاسیون است. همچنین فرض می‌کنیم که فوتون‌های ساطع شده از منبع هر بار مستقل از یکدیگر هستند. ما می‌خواهیم دمای منبع را با اندازه گیری‌های پی در پی فوتون‌های ساطع شده تخمین بزنیم.

آ) ضریب نرمالیزاسیون $c(\theta)$ را بیابید.

ب) امید ریاضی و واریانس تعداد فوتون های ساطع شده را بیابید.

ج) تخمین ML برای دمای منبع $\psi = \frac{۱}{\theta}$ را بر اساس $K_۱, K_۲, ...$ که تعداد فوتون های ساطع شده پس از n بار اندازه گیری است بیابید.

د) نشان دهید که تخمین بدست آمده پایدار است.

پاسخ .

سوال ۱) (آ)

S = number of passangers who show up

$$P(S \geqslant ۳۰۱) = P(S > ۳۰۱ - ۰/۵) = ۰/۰۴۹$$

(ب)

$$P(S \geqslant ۱۵۱) = P(\frac{S-\mu}{\sigma} \geqslant \frac{۱۵۱ - ۰/۵ - \mu}{\sigma}) \approx ۰/۰۰۱$$

سوال ۲) (آ) درصد فارغ التحصیلان که کار پیدا کرده اند:

$$\hat{p} = \frac{۳۴۸}{۴۰۰} = ۰/۸۷$$

(ب) شرایط:

$$n \times p \times q > ۱۰ \longrightarrow ۴۰۰(۰/۸۷)(۱ - ۰/۸۷) = ۳۴۸(۰/۱۳) = ۴۵/۲۴ > ۱۰ \checkmark$$

(ج)

$$CI = \hat{p} \pm z_{۱-\frac{\alpha}{۲}}\sqrt{\frac{\hat{p}(۱-\hat{p})}{n}}$$
$$= ۰/۸۷ \pm ۰/۰۳ = (۰/۸۴, ۰/۹)$$

با اطمینان ۹۵٪ میدانیم نسبت فارغ التحصیلانی که در مدت یکسال پس از اتمام دوره ی کارشناسی خود در این دانشگاه کار پیدا کرده اند بین $(۰/۸۴, ۰/۹)$ است.

(د) عرض بازه‌ی اطمینان ۹۹٪ بیشتر است. هرچه عرض یک بازه‌ی اطمینان بیشتر باشد، اطمینان بیشتری داریم که نسبت مورد نیاز در این بازه قرار گیرد.

سوال ۳)

$$S_n = X_۱ + X_۲ + ...$$

$$P(S_n \times c > n(۱+\eta)pc) = P(S_n > n(۱+\eta)p) \leq \delta$$

$$\longrightarrow n \geq \frac{-۲(۱-p) \times ln(\delta)}{p\eta^۲}$$

سوال ۴) (آ) بازه اطمینان برای میانگین با واریانس نامعلوم :

$$\left(\bar{X}-\frac{S}{\sqrt{n}}z_{۱-\frac{\alpha}{۲}}<\mu<\bar{X}+\frac{S}{\sqrt{n}}z_{۱-\frac{\alpha}{۲}}\right)$$

بازه‌ی اطمینان ٪۹۰ :

$$\left(۲/۱-\frac{۰/۲}{\sqrt{۶}}(۱/۶۴۵)<\mu<۲/۱+\frac{۰/۲}{\sqrt{۶}}(۱/۶۴۵)\right)$$

بازه‌ی اطمینان ٪۹۵ :

$$\left(۲/۱-\frac{۰/۲}{\sqrt{۶}}(۱/۹۶)<\mu<۲/۱+\frac{۰/۲}{\sqrt{۶}}(۱/۹۶)\right)$$

(ب) بازه اطمینان ٪۹۰:

$$\left(۲/۲۵-\frac{۰/۴۶}{\sqrt{۶}}(۱/۶۴۵)<\mu<۲/۲۵+\frac{۰/۴۶}{\sqrt{۶}}(۱/۶۴۵)\right)$$

بازه اطمینان ٪۹۵:

$$\left(۲/۲۵-\frac{۰/۴۶}{\sqrt{۶}}(۱/۹۶)<\mu<۲/۲۵+\frac{۰/۴۶}{\sqrt{۶}}(۱/۹۶)\right)$$

	بدون داده پرت	با داده پرت
بازه اطمینان ۹۰ درصد	(۱/۹۶, ۲/۲۳)	(۱/۹۴, ۲/۵۶)
بازه اطمینان ۹۵ درصد	(۱/۹۳, ۲/۲۶)	(۱/۸۸, ۲/۶۱)

همانطور که مشاهده می‌کنید در صورت وجود داده‌ی دور افتاده، بازه اطمینان بزرگ‌تر می‌شود.

سوال ۵) (آ) بازه اطمینان :

$$\left(۰/۶۳-\sqrt{\frac{۰/۲۳۳۱}{n}}\times ۱/۹۶, ۰/۶۳+\sqrt{\frac{۰/۲۳۳۱}{n}}\times ۱/۹۶\right)$$

با توجه حاشیه‌ی خطای داده شده در صورت سوال:

$$\sqrt{\frac{۰/۲۳۳۱}{n}}\times ۱/۹۶=۰/۰۳\longrightarrow n=۹۹۴/۹۷۴۴\longrightarrow n=۹۹۵$$

دقت شود که n باید عدد طبیعی باشد.

(ب)

$$(۰/۶۳-۰/۰۴۹۵, ۰/۶۳+۰/۰۴۹۵)=(۰/۵۸۰۵, ۰/۶۷۹۵)$$

سوال ۶) از نامساوی چبی‌شف استفاده می‌کنیم. بنابراین:

$$P[|X-۱۵|\geq ۵]\leq\frac{۹}{۲۵}=۰/۳۶$$

سوال ۷) (آ) در متغیر تصادفی گسسته مجموع احتمال پیشامد ها ۱ می‌شود با استفاده از این نکته:

$$۱=\sum_{k=۰}^{\infty}c(\theta)e^{-k\theta}=\frac{c(\theta)}{۱-e^{-\theta}}$$

$$c(\theta)=۱-e^{-\theta}$$

(ب) با دقت می‌توان دید که K دارای توزیع هندسی شیفت یافته با $p=۱-e^{-\theta}$ است(K از صفر شروع شده است) پس داریم:

$$E_\theta[K]=\frac{۱-p}{p}-۱=\frac{e^{-\theta}}{۱-e^{-\theta}}$$

$$var_\theta(K)=\frac{۱-p}{p^۲}=\frac{e^{-\theta}}{(۱-e^{-\theta})^۲}$$

(ج) ابتدا متغیر K_i را تعداد فوتون های ساطع شده در زمان i ام می‌نامیم توزیع توام برای متغیر تصادفی K بصورت زیر است:

$$p_K(k_۱, ..., k_n; \theta) = c(\theta)^n \prod_{i=۱}^{\infty} e^{-\theta k_i} = c(\theta)^n e^{-\sum K_i \times \theta}$$

$$L(\theta) = (۱ - e^{-\theta})^n \times e^{-\sum K_i \times \theta}$$

با گرفتن لگاریتم تابع log-likelihood را بدست می‌آوریم. سپس از تابع بدست آمده نسبت به θ مشتق می گیریم و برابر صفر قرار می‌دهیم. در انتها خواهیم داشت:

$$\Longrightarrow \quad \hat{\psi_n} = \frac{۱}{ln(۱ + \frac{n}{\sum K_i})}$$

(د) اگر $S_n = \sum K_i$ باشد، از قانون ضعیف اعداد بزرگ داریم:

$$P\Big\{\frac{S_n}{n} = E(K) = \frac{e^{-\theta}}{۱ - e^{-\theta}}\Big\} \to ۱$$

از آنجایی که داریم:

$$۱ + \frac{n}{S_n} = \frac{e^{-\theta} + ۱ - e^{-\theta}}{e^{-\theta}} = e^{\theta}$$

با کمک از نتایج بخش قبل می‌توان گفت:

$$\hat{\theta_n} = ln(۱ + \frac{n}{\sum K_i}) \to \theta$$

به همین صورت برای ψ نیز می‌توانیم بنویسیم:

$$\hat{\psi_n} = \frac{۱}{\hat{\theta_n}} \to \psi$$

از همگرایی به احتمال ۱ می‌توانیم به همگرایی در احتمال برسیم پس نتیجه می گیریم که این ۲ متغیر پایدار خواهند شد