- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص‌شده است.
- هم‌کاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

سوالات نظری (۹۵ نمره)

مسئله‌ی ۱. (۱۰ نمره)

فرض کنید دو تاس مستقل از هم و fair و متمایز داریم. در هر بار بازی شما به عنوان بازیکن می توانید یک حدسی برای هرکدام از تاس‌ها انجام دهید و مقداری پول در جعبه قرار دهید. در صورتی که هیچ‌کدام از حدس های شما درست نباشد شما پول خود را از دست خواهید داد. در غیر این صورت شما n (=تعداد حدس‌های درست) برابر پول جعبه به همراه اصل پول خود را برنده خواهید شد.
برای بازی بالا، برای هردست بازی، امیدریاضی سود حاصله را بدست آوردید.

حل. ۳ حالت ممکن است پیش بیاید.

- پیش بینی هیچ کدام از تاس‌ها درست نباشد ($n = ۰$)
- پیش بینی یکی از تاس ها درست باشد ($n = ۱$)
- پیش بینی هردو تاس درست باشد ($n = ۲$)

با فرض سرمایه‌گذاری به مقدار x در هر بازی و باتوجه به فرضیات مساله داریم:

$$\begin{aligned}\mathbb{E}\left[profit\right] =& \sum_{i=۰}^{۲} profit\left(n=i\right)\times P(n=i)\\ =&(-x)\times\left(\frac{۵}{۶}\right)^{۲}+(x)\times\left(۲\times\frac{۵}{۶}\times\frac{۱}{۶}\right)+(۲x)\times\left(\frac{۱}{۶}\right)^{۲}=-\frac{۱۳}{۳۶}x\end{aligned}$$

▷

مسئله‌ی ۲. (۱۰ نمره)

X و Y متغیرهای تصادفی نمایی با پارامترهای $\lambda_X = ۲, \lambda_Y = ۵$ و مستقل از هم هستند. متغیرهای تصادفی Z و W را به این صورت تعریف میکنیم:

$$W = \max(X, Y)$$
$$Z = \min(X, Y)$$

تابع چگالی احتمال را برای $S = Z + W$ بدست آورید.

حل.

$$S = Z + W = \min(X,Y) + \max(X,Y) = X + Y$$

$$\begin{aligned} P(S=s) &= \int_{-\infty}^{+\infty} P(X = s-r)\, P(Y=r) \\ &= \int_{۰}^{s} ۲e^{-۲(s-r)} \times ۵e^{-۵r} dr \\ &= ۱۰ e^{-۲s} \int_{۰}^{s} e^{-۳r} dr \\ &= \frac{۱۰\left(e^{-۲s} - e^{-۵s}\right)}{۳} \qquad\qquad s \geqslant ۰ \end{aligned}$$

$\triangleright$

مسئله‌ی ۳. (۲۰ نمره)

برای متغیرهای تصادفی دلخواه X و Y عبارات زیر را اثبات نمایید:

(آ) با فرض مستقل بودن X و Y : $\mathbb{E}[X|Y=y] = \mathbb{E}[X]$

(ب) $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

(ج) $\mathrm{var}(X) = \mathbb{E}[\mathrm{var}(X|Y)] + \mathrm{var}[\mathbb{E}(X|Y)]$

(د) $\mathbb{E}[XY] = \mathbb{E}[Y\mathbb{E}[X|Y]]$

حل.

(آ)

$$\begin{aligned} \mathbb{E}[X|Y=y] &= \int_{-\infty}^{+\infty} x.f(X=x|Y=y)dx \\ &= \int_{-\infty}^{+\infty} x.f(X=x)dx = \mathbb{E}[X] \end{aligned}$$

(ب)

$$\begin{aligned} \mathbb{E}[\mathbb{E}[X|Y]] &= \int_{-\infty}^{+\infty} \mathbb{E}[X|Y=y].f(Y=y)\,dy \\ &= \int_{-\infty}^{+\infty} (\int_{-\infty}^{+\infty} x.f(X=x|Y=y)dx).f(Y=y)\,dy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x.f(X=x, Y=y)dxdy \\ &= \int_{-\infty}^{+\infty} x.f(X=x)dx = \mathbb{E}[X] \end{aligned}$$

(ج) طبق تعریف داریم:

$$\mathrm{var}(X) = \mathbb{E}\left[X^{۲}\right] - \mathbb{E}^{۲}[X]$$

طبق نتیجه بخش ب داریم:

$$\mathbb{E}\left[X^{2}\right]-\mathbb{E}^{2}[X]=\mathbb{E}\left[\mathbb{E}\left[X^{2}|Y\right]\right]-\mathbb{E}^{2}[\mathbb{E}[X|Y]]$$

عبارت $\mathbb{E}\left[\mathbb{E}^{2}[X|Y]\right]$ را به عبارت بدست آمده اضافه و کم می کنیم:

$$\begin{aligned}
&=\mathbb{E}\left[\mathbb{E}\left[X^{2}|Y\right]\right]-\mathbb{E}^{2}[X]+\mathbb{E}\left[\mathbb{E}^{2}[X|Y]\right]-\mathbb{E}\left[\mathbb{E}^{2}[X|Y]\right]\\
&=\left(\mathbb{E}\left[\mathbb{E}\left[X^{2}|Y\right]\right]-\mathbb{E}\left[\mathbb{E}^{2}[X|Y]\right]\right)+\left(\mathbb{E}\left[\mathbb{E}^{2}[X|Y]\right]-\mathbb{E}^{2}[\mathbb{E}[X|Y]]\right)\\
&=\left(\mathbb{E}\left[\mathbb{E}\left[X^{2}|Y\right]-\mathbb{E}^{2}[X|Y]\right]\right)+\left(\mathbb{E}\left[\mathbb{E}^{2}[X|Y]\right]-\mathbb{E}^{2}[\mathbb{E}[X|Y]]\right)\\
&=\mathbb{E}[\operatorname{var}(X|Y)]+\operatorname{var}[\mathbb{E}(X|Y)]
\end{aligned}$$

(د)

$$\begin{aligned}
\mathbb{E}[XY]&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy.f(X=x,Y=y)dxdy\\
&=\int_{-\infty}^{+\infty}y\left(\int_{-\infty}^{+\infty}x.f(X=x|Y=y)dx\right).f(Y=y)dy\\
&=\mathbb{E}[Y\mathbb{E}[X|Y]]
\end{aligned}$$

▷

مسئله‌ی ۴. (۲۰ نمره)

با فرض اینکه A یک ماتریس n*n معکوس پذیر می باشد، موارد زیر را ثابت کنید.

(آ) اگر λ یکی از مقدار ویژه های ماتریس A باشد، $\frac{1}{\lambda}$ نیز یکی از مقدار ویژه های ماتریس A^{-1} با همان بردار ویژه خواهد بود.

(ب) اگر A یک ماتریس مثلثی باشد،عناصر روی قطر ماتریس مقادیر ویژه آن خواهند بود.

(ج) ضرب مقدار ویژه های ماتریس A برابر دترمینان ماتریس A خواهد بود.

(د) جمع مقدار ویژه های ماتریس A برابر $trace(A)$ خواهد بود.

حل.

(آ) باتوجه به مفروضات مساله داریم:

$$Av=v\lambda$$

باتوجه به معکوس پذیری ماتریس A با ضرب معکوس ماتریس A از سمت چپ هر طرف تساوی داریم:

$$\begin{aligned}
A^{-1}Av&=A^{-1}v\lambda\\
v&=A^{-1}v\lambda
\end{aligned}$$

باتوجه به اینکه λ ماتریس قطری می باشد و ماتریس A معکوس‌پذیر بود، پس هیچ‌کدام از مقادیر روی قطر اصلی ماتریس λ صفر نمی باشد و ماتریس λ معکوس‌پذیر است. معکوس ماتریس قطری خود یک ماتریس قطری با معکوس همان مقادیر روی قطر می باشد. پس با ضرب معکوس ماتریس λ از راست داریم:

$$\begin{aligned}
v\lambda^{-1}&=A^{-1}v\lambda\lambda^{-1}\\
v\lambda^{-1}&=A^{-1}v
\end{aligned}$$

که رابطه بدست آمده بیانگر مطلوب سوال است.

(ب) برای دترمینان اگر رابطه Leibniz را که از این لینک می توانید آن را مشاهده کنید در نظر بگیرید، با توجه به اینکه ماتریس $A-\lambda I$ نیز خود ماتریس مثلثی می باشد دترمینان آن برابر ضرب عناصر روی قطر خواهد بود و ریشه‌های رابطه که بیانگر مقادیر ویژه می‌باشند، عناصر روی قطر خواهند بود.

(ج) طبق روابط polynomial characteristic یک ماتریس داریم:

$$|\lambda I-A|=p(\lambda)=(\lambda-\lambda_۱)(\lambda-\lambda_۲)....(\lambda-\lambda_n)$$

در رابطه بالا با قرار دادن $\lambda=۰$ داریم:

$$|-A|=\prod_{i=۱}^{n}(-\lambda_i)=(-۱)^n\prod_{i=۱}^{n}\lambda_i$$

از طرفی برای دترمینان منفی A داریم که:

$$|-A|=(-۱)^n.|A|$$

با مقایسه روابط به مطلوب سوال دست پیدا می‌کنیم.

(د) برای دترمینان اگر رابطه Leibniz را که از این لینک می توانید آن را مشاهده کنید در نظر بگیرید، داریم:

$$|\lambda I-A|=(\lambda-a_{۱,۱})(\lambda-a_{۲,۲})....(\lambda-a_{n,n})+f(\lambda)$$

که در رابطه بالا $f(\lambda)$ یک چندجمله‌ای از مرتبه $n-۲$ می باشد.
از طرفی داریم:

$$|\lambda I-A|=p(\lambda)=(\lambda-\lambda_۱)(\lambda-\lambda_۲)....(\lambda-\lambda_n)$$

حال اگر ضریب $\lambda^{n-۱}$ دو رابطه بالا را با هم مقایسه کنیم به این نتیجه می‌رسیم که:

$$-\left(a_{۱,۱}+a_{۲,۲}+\cdots+a_{n,n}\right)=-(\lambda_۱+\lambda_۲+\cdots+\lambda_n)$$

که عبارت سمت چپ همان $trace(A)$ می باشد و مطلوب مساله بدست می‌آید.

▷

مسئله‌ی ۵. (۱۵ نمره)

با فرض اینکه a و x بردارهای ستونی و A ماتریس مربعی باشد،روابط زیر را برای مشتقات ماتریسی اثبات کنید.

(آ) $\frac{da^Tx}{dx}=a^T$

(ب) $\frac{dx^TAx}{dx}=x^T(A+A^T)$

(ج) $\frac{d(x^Ta)^۲}{dx}=۲x^Taa^T$

حل. در حل تمام مسایل notation های نوشته شده در این لینک را مبنا قرار می‌دهیم.

(آ) مطلوب سوال مشتق یک scalar نسبت به یک بردار ستونی است، بنابراین مشتق یک بردار سطری(B) است و برای المان های آن داریم:

$$B_i = \frac{d\left(a^T x\right)}{dx_i}$$

با محاسبه رابطه بالا به بردار a^T میرسیم.

(ب) طبق product rule در مشتق‌گیری که قابل تعمیم به فضای ماتریسی هست و نتیجه بدست آمده در بخش قبل داریم:

$$\frac{dx^T Ax}{dx} = (Ax)^T + x^T \frac{d\,(Ax)}{dx}$$

مشتق دوم نیز مشتق یک بردار نسبت به یک بردار است که طبق رابطه نوشته شده در لینکی که ابتدا قرار داده شده‌است، داریم:

$$Ax = \begin{bmatrix} \sum_{i=۱}^{n} A_{۱,i}x_i & \sum_{i=۱}^{n} A_{۲,i}x_i & \cdots & \sum_{i=۱}^{n} A_{n,i}x_i \end{bmatrix}^T$$

$$\tfrac{d(Ax)}{dx} = A$$

پس داریم:

$$\frac{dx^T Ax}{dx} = x^T A^T + x^T A = x^T(A + A^T)$$

(ج) ماتریسی که مشتق آن خواسته شده است را می‌توان به صورت زیر بازنویسی کرد:

$$\left(x^T a\right)^۲ = \left(x^T a\right)\left(x^T a\right)^T = x^T a a^T x$$

حال اگر فرض کنیم $C = aa^T$ طبق خاصیت بدست آمده در بخش قبل داریم:

$$\frac{d\left(x^T a\right)^۲}{dx} = \frac{d\left(x^T Cx\right)}{dx} = x^T\left(C + C^T\right) = ۲x^T C = ۲x^T aa^T$$

▷

مسئله‌ی ۶. (۲۰ نمره)

در این سوال، هدف اصلی محاسبه MLE و MAP برای میانگین یک توزیع نرمال تک متغیره است. در این سوال ما فرض می کنیم که N نمونه $x_۱, \cdots, x_N$ به طور مستقل از یک توزیع نرمال با واریانس معلوم $\sigma^۲$ و میانگین نامعلوم μ گرفته شده اند.

۱. تخمین گر MLE برای μ را محاسبه کنید. تمامی محاسبات مورد استفاده را به طور دقیق ذکر کنید.

۲. حال تخمین گر MAP را برای μ بدست آورید. برای این بخش فرض کنید که توزیع پیشین (Prior) برای μ خود یک توزیع نرمال با میانگین ν و واریانس $\beta^۲$ باشد. تمامی محاسبات خود را به طور دقیق ذکر کنید. راهنمایی: می‌توانید از این نکته استفاده کنید که:

$$\beta^2\left(\sum_{i=1}^{N}\left(x_i-\mu\right)^2\right)+\sigma^2(\mu-\nu)^2$$

$$=\left[\mu\sqrt{N\beta^2+\sigma^2}-\frac{\sigma^2\nu+\beta^2\sum_{i=1}^{N}x_i}{\sqrt{N\beta^2+\sigma^2}}\right]^2-\frac{\left[\sigma^2\nu+\beta^2\sum_{i=1}^{N}x_i\right]^2}{N\beta^2+\sigma^2}+\beta^2\left(\sum_{i=1}^{N}x_i^2\right)+\sigma^2\nu^2$$

البته توجه کنید که بدون استفاده از این راهنمایی هم امکان حل سوال وجود دارد.

حل.

۱. ابتدا عبارت Likelihood را بدست می‌آوریم:

$$\begin{aligned}P\left(x_1,\ldots,x_N\mid\mu\right)&=\prod_{i=1}^{N}P\left(x_i\mid\mu\right)\\&=\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{\left(z_i-\mu\right)^2}{2\sigma^2}}\end{aligned}$$

حال از این عبارت لگاریتم می‌گیریم تا log-likelihood را بدست بیاوریم:

$$\log\left(P\left(x_1,\ldots,x_N\mid\mu\right)\right)=\sum_{i=1}^{N}\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)-\frac{\left(x_i-\mu\right)^2}{2\sigma^2}$$

حال مشتق آن را نسبت به μ حساب می کنیم:

$$\frac{d\log\left(P\left(x_1,\ldots,x_N\mid\mu\right)\right)}{d\mu}=\sum_{i=1}^{N}\frac{\left(x_i-\mu\right)}{\sigma^2}$$

با قرار دادن مشتق برابر صفر داریم:

$$\begin{aligned}0&=\sum_{i=1}^{N}\frac{\left(x_i-\mu\right)}{\sigma^2}\\0&=\sum_{i=1}^{N}\left(x_i-\mu\right)\\\sum_{i=1}^{N}\mu&=\sum_{i=1}^{N}x_i\\N\mu&=\sum_{i=1}^{N}x_i\\\hat{\mu}&=\frac{\sum_{i=1}^{N}x_i}{N}\end{aligned}$$

۲. این بخش را می‌توان با دو راه حل کرد. ابتدا راه کوتاه‌تر را بررسی می‌کنیم و سپس سراغ راهی که اندکی طولانی‌تر است می رویم.

- راه اول:

با استفاده از قاعده بیز داریم:

$$P\left(\mu \mid x_{1}, \ldots x_{n}\right)=\frac{P\left(x_{1}, \ldots, x_{N} \mid \mu\right) P(\mu)}{P\left(x_{1}, \ldots, x_{N}\right)}$$

با استفاده از بخش اول داریم که:

$$P\left(x_{1}, \ldots, x_{N} \mid \mu\right)=\prod_{i=1}^{N} \frac{1}{\sqrt{2 \pi \sigma^{2}}} e^{-\frac{\left(x_{i}-\mu\right)^{2}}{2 \sigma^{2}}}$$

همچنین با توجه به داده سوال در مورد توزیع μ داریم:

$$P(\mu)=\frac{1}{\sqrt{2 \pi \beta^{2}}} e^{-\frac{(\mu-\nu)^{2}}{2 \beta^{2}}}$$

حال باید μ ای را پیدا کنیم که مقدار زیر را بیشینه کند:

$$P\left(\mu \mid x_{1}, \ldots x_{n}\right)=\frac{\left(\prod_{i=1}^{N} \frac{1}{\sqrt{2 \pi \sigma^{2}}} e^{-\frac{\left(x_{i}-\mu\right)^{2}}{2 \sigma^{2}}}\right) \frac{1}{\sqrt{2 \pi \beta^{2}}} e^{-\frac{(\mu-\nu)^{2}}{2 \beta^{2}}}}{C}$$

مقدار مخرج در این جا اهمیتی ندارد، چون در نهایت در فرآیند مشتق گیری حذف خواهد شد.
با این توصیف از دو طرف log می‌گیریم.

$$\log \left(P\left(\mu \mid x_{1}, \ldots x_{n}\right)\right)=\left(\sum_{i=1}^{N}-\log \left(\sqrt{2 \pi \sigma^{2}}\right)-\frac{\left(x_{i}-\mu\right)^{2}}{2 \sigma^{2}}\right)-\log \left(\sqrt{2 \pi \beta^{2}}\right)-\frac{(\mu-\nu)^{2}}{2 \beta^{2}}$$

با مشتق گیری داریم:

$$\frac{\partial \log \left(P\left(\mu \mid x_{1}, \ldots x_{n}\right)\right)}{\partial \mu}=\left(\sum_{i=1}^{N} \frac{x_{i}-\mu}{\sigma^{2}}\right)-\frac{\mu-\nu}{\beta^{2}}$$

با صفر قرار دادن آن داریم:

$$\begin{aligned}
0 &=\left(\sum_{i=1}^{N} \frac{x_{i}-\mu}{\sigma^{2}}\right)-\frac{\mu-\nu}{\beta^{2}} \\
\frac{\mu-\nu}{\beta^{2}} &=\sum_{i=1}^{N} \frac{x_{i}-\mu}{\sigma^{2}} \\
\frac{\mu-\nu}{\beta^{2}} &=-\frac{\sum_{i=1}^{N} x_{i}}{\sigma^{2}}-\frac{N \mu}{\sigma^{2}} \\
\frac{\mu}{\beta^{2}}+\frac{N \mu}{\sigma^{2}} &=\frac{\sum_{i=1}^{N} x_{i}}{\sigma^{2}}+\frac{\nu}{\beta^{2}} \\
\frac{\left(\sigma^{2}+N \beta^{2}\right) \mu}{\sigma^{2} \beta^{2}} &=\frac{\sigma^{2} \nu+\beta^{2} \sum_{i=1}^{N} x_{i}}{\sigma^{2} \beta^{2}} \\
\hat{\mu} &=\frac{\sigma^{2} \nu+\beta^{2} \sum_{i=1}^{N} x_{i}}{\sigma^{2}+N \beta^{2}}
\end{aligned}$$

- راه دوم:

در این روش ابتدا نشان می‌دهیم که توزیع Posterior نرمال خود نیز یک توزیع نرمال خواهد بود و سپس با کمک این نکته که توزیع نرمال در نقطه میانگین خود بیشینه می‌شود سوال را حل کرده و MAP را می‌یابیم.
ابتدا توجه کنید که مطابق راه قبلی می‌دانیم:

$$P\left(\mu \mid x_1, \ldots x_n\right)=\frac{\left(\prod_{i=1}^N \frac{1}{\sqrt{2 \pi \sigma^2}} e^{-\frac{\left(x_i-\mu\right)^2}{\sigma^2}}\right) \frac{1}{\sqrt{2 \pi \beta^2}} e^{-\frac{(\mu-\nu)^2}{2 \beta^2}}}{C}$$

که در آن $C=P(x_1, \cdots, x_N)$ اما این بار به جای بیشینه کردن این عبارت، نشان می‌دهیم که این عبارت باید از توزیع نرمال پیروی کند.

$$\begin{aligned}
P\left(\mu \mid x_1, \ldots x_n\right) & \propto\left(\prod_{i=1}^N \frac{1}{\sqrt{2 \pi \sigma^2}} e^{-\frac{\left(x_i-\mu\right)^2}{2 \sigma^2}}\right) \frac{1}{\sqrt{2 \pi \beta^2}} e^{-\frac{(\mu-\nu)^2}{2 \beta^2}} \\
& \propto\left(\prod_{i=1}^N e^{-\frac{\left(x_i-\mu\right)^2}{2 \sigma^2}}\right) e^{-\frac{(\mu-\nu)^2}{2 \beta^2}} \\
& =e^{-\frac{1}{2 \sigma^2}\left(\sum_{i=1}^N\left(x_i-\mu\right)^2\right)} e^{-\frac{(\mu-\nu)^2}{2 \beta^2}} \\
& =e^{-\frac{1}{2 \sigma^2}\left(\sum_{i=1}^N\left(x_i-\mu\right)^2\right)-\frac{(\mu-\nu)^2}{2 \beta^2}} \\
& =e^{-\frac{\beta^2\left(\sum_{i=1}^N\left(x_i-\mu\right)^2\right)+\sigma^2(\mu-\nu)^2}{2 \sigma^2 \beta^2}}
\end{aligned}$$

در این جا از راهنمایی داده شده در صورت سوال استفاده کرده و داریم:

$$P\left(\mu \mid x_1, \ldots x_n\right) \propto e^{-\frac{1}{2 \sigma^2 \beta^2}\left(\left[\mu \sqrt{N \beta^2+\sigma^2}-\frac{\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i}{\sqrt{N \beta^2+\sigma^2}}\right]^2-\frac{\left[\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i\right]^2}{N \beta^2+\sigma^2}+\beta^2\left(\sum_{i=1}^N x_i^2\right)+\sigma^2 \nu^2\right)}$$

$$=e^{-\frac{1}{2 \sigma^2 \beta^2}\left(\left[\mu \sqrt{N \beta^2+\sigma^2}-\frac{\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i}{\sqrt{N \beta^2+\sigma^2}}\right]^2\right)} e^{\frac{1}{2 \sigma^2 \beta^2}\left(\frac{\left[\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i\right]^2}{N \beta^2+\sigma^2}+\beta^2\left(\sum_{i=1}^N x_i^2\right)+\sigma^2 \nu^2\right)}$$

$$\propto e^{-\frac{1}{2 \sigma^2 \beta^2}\left(\left[\mu \sqrt{N \beta^2+\sigma^2}-\frac{\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i}{\sqrt{N \beta^2+\sigma^2}}\right]^2\right)}$$

$$=e^{-\frac{N \beta^2+\sigma^2}{2 \sigma^2 \beta^2}\left(\left[\mu-\frac{\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i}{N \beta^2+\sigma^2}\right]^2\right)}$$

یعنی این یک تابع نرمال با میانگین $\frac{\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i}{N \beta^2+\sigma^2}$ و واریانس $\frac{\sigma^2 \beta^2}{N \beta^2+\sigma^2}$ است. از آن جایی که مقدار حداکثر تابع نرمال در میانگین آن اتفاق می‌افتد، تخمین گر MAP برابر است با:

$$\hat{\mu}_{\mathrm{MAP}}=\frac{\sigma^2 \nu+\beta^2 \sum_{i=1}^N x_i}{N \beta^2+\sigma^2}$$

▷

سوالات عملی (۱۵ نمره)

مسئله‌ی ۷. (۱۵ نمره)

در این سوال می خواهیم تاثیر تعداد داده های نمونه برداری شده را بر مقادیر بدست آمده توسط تخمینگرهای MLE و MAP بررسی کنیم.
یک توزیع به فرم گاوسی تک متغیره با پارامترهای $x \sim N\left(\mu = ۲, \sigma^۲ = ۵\right)$ را در نظر بگیرید که داده های موجود از این توزیع به صورت i.i.d نمونه برداری می‌شوند. همچنین فرض کنید توزیع پیشین μ یک توزیع گاوسی با پارامترهای $N\left(۱, ۲\right)$ می باشد.
به ترتیب ۱۰، ۱۰۰ و ۱۰۰۰ نمونه از توزیع گاوسی اولیه نمونه بردارید. در هر حالت با کمک داده های نمونه برداری شده و توزیع پیشین بالا و روابط بدست آمده در مساله قبل، با استفاده از تخمین گرهای MLE و MAP مقدار μ را تخمین بزنید.
نتایج بدست آمده را تفسیر کنید. (به دلیل ذات تصادفی بودن نمونه های برداشته شده، برای تفسیر دقیق‌تر توصیه می‌شود چندبار عمل بالا را تکرار نمایید.)

موفق باشید :)