BIMU2004
Olasılık Teorisi ve İstatistik
Final Sınavı - Sorular ve Çözümleri

İstanbul Üniversitesi - Cerrahpaşa
Bilgisayar Mühendisliği Bölümü - Güz 2020

Son yükleme: 13.Ocak.2020 12:40

LÜTFEN OKUYUN

Çözümlerinizi beyaz renkli düz, çiizgili veya kareli bir kağıda yapınız.

Çözümlerinizde ne yaptığınızı açıklamanız, yaptığınız işte kullandığınız formülleri yazmanız lazım. Formül isimlerini yazacaksınız, Bayes Fomülü ise "Bayes Formülü" demeniz lazım.

Cevabınızı kare içine alın. Kare içine almazsanız son yazdığınız şeyi cevap olarak kabul edeceğim.

Gereksiz bilgi yazmayın. Fazladan bişey yazayım, ya tutarsa gibi düşünürseniz puan alamazsınız.

Sınavı çözdükten sonra **her sayfaya isim, soyisim, numara yazıp ve imza atınız**. Sonra **Microsoft Office Lens**, **Adobe Scan** vs. gibi "döküman modu" olan bir tarayıcı mobil uygulama veya masaüstü tarayıcı kullanarak sınav çözümlerinizi tek dosya PDF'e çevirin.

Dosya ismini aşağıdaki gibi yapınız, `1306XXXX` yazan yere kendi numaranızı yazınız:

`olasilik-2020-final-1306XXXX.pdf`

Son dakika gelmeden PDF dosyanızı MERGEN'e yükleyin

Kolay gelsin. (Mustafa Dağtekin)

SORULAR ve ÇÖZÜMLER

S1: Bi mağazaya müşteri gelme süreci Poisson dağılımı ile modelleniyor ve bu mağazaya saatte ortalama 3 müşteri giriyor. Aşağıdaki soruları cevaplayınız:

(a) `(10 puan)` Bu mağazaya 40 dakikada 2'den fazla müşteri girme olasılığı nedir?

Çözüm 1a:

X: Mağazaya 40 dakikada giren müşteri sayısı ise, X bir Poisson rastegele değişkenidir ve parametresi aşağıdaki gibi bulunur. $\lambda = 3$ müşteri/saat ve $T = 40$ dk verilmiş.

$$\begin{aligned}\alpha &= \lambda\, T \\ &= 3 \text{ müşteri/saat } \times \frac{40}{60} \text{ saat} \\ &= 2 \text{ müşteri}\end{aligned}$$

X'in O.K.F.'si

$$\begin{aligned}P(X = x) = \mathrm{p}(x) &= \mathrm{e}^{-\alpha}\, \frac{\alpha^x}{x!} \\ &= \mathrm{e}^{-2}\, \frac{2^x}{x!}\end{aligned}$$

Bize sorulan $P(X > 2)$, bunu da

$$P(X > 2) = 1 - P(X \leq 2)$$

kullanarak bulabiliriz. Burdan:

$$\begin{aligned}P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \mathrm{e}^{-2}(1 + 2 + 2^2/2) \\ &= 0.6766\end{aligned}$$

$$\begin{aligned}P(X > 2) &= 1 - 0.6766 \\ P(X > 2) &= 0.3234 \qquad \blacksquare\end{aligned}$$

(b) `(10 puan)` Bu mağazaya 4. müşteri girdikten sonra 6. müşteri girinceye kadar geçen ortalama süre nedir?

Çözüm 1b:

Burda 4 müşteri girene kadarki zaman dikkate alınmıyor. O zaman burda rastegele değişkeni şöyle tanımlayabiliriz:

X: 2 müşteri girinceye kadar geçen zaman ise , X Gama dağılımlı bir rastgele değişkendir ve parametreleri: $\lambda = 3$ müşteri/saat, $r = 2$'dir. X'in OYF:

$$f(x) = \frac{\lambda^r \; x^{r-1} \; \mathrm{e}^{-\lambda x}}{\Gamma(r)} \;, \qquad x > 0$$

r tamsayı olduğu için, $\Gamma(r) = (r-1)!$, o zaman :

$$\begin{aligned} f(x) &= \frac{\lambda^r \; x^{r-1} \; \mathrm{e}^{-\lambda x}}{(r-1)!} \\ &= \frac{3^2 \; x^{2-1} \; \mathrm{e}^{-3x}}{(2-1)!} \\ &= 9 \; x \; \mathrm{e}^{-3x} \end{aligned}$$

olur. Gama dağılımın ortalaması aşağıdaki şekilde hesaplanır.

$$\begin{aligned} E(X) &= \frac{r}{\lambda} \\ E(X) &= \frac{2}{3} = 0.666667 \qquad \blacksquare \end{aligned}$$

Not: Bunu 2 adet üstel dağılımlı RD'nin toplamını kullanarak da çözebilirdik.

X_1 ve X_2, $\lambda = 3$ hata/saat olan üstel dağılımlı R.D.'ler ise bize sorulan $E(X_1 + X_2)$. O da $E(X_1) + E(X_2) = 1/\lambda + 1/\lambda = 2/\lambda$ olur.

Not: Sorunun soruş şekli itibari, 4. müşteri girinceye kadar geçen zaman bilinmese de saymaya o vakitten itibaren başlayabiliriz, çünkü o kısmın "geçtiği" varsayılıyor, emri vakidir artık. Fakat bu Gama dağılımda hafızasızlık özelliği olduğu anlamına gelmez. Gama dağılımında hafızasızlık özelliği yoktur. Sadece üstel ve geometrik dağılımda bu özellik mevcuttur.

(c) `(10 puan)` Bu mağazaya 4. müşteri girdikten sonra 6. müşteri girinceye kadar geçen sürenin 50 dakikadan az olma ihtimali nedir?

Çözüm 1c:

Bize sorulan $P(X < \frac{50}{60})$.

$$\begin{aligned} P(X < \frac{50}{60}) &= \int_0^{5/6} f(x)\ \mathrm{d}x \\ &= \int_0^{5/6} 9\ x\ \mathrm{e}^{-3x}\ \mathrm{d}x \\ u &= 9\ x \\ \mathrm{d}u &= 9\ \mathrm{d}x \\ \mathrm{d}v &= \mathrm{e}^{-3x}\ \mathrm{d}x \\ v &= (-1/3)\ \mathrm{e}^{-3x} \end{aligned}$$

$$\begin{aligned} \int u\ \mathrm{d}v &= u\ v - \int v\ \mathrm{d}u \qquad \text{ile belirsiz integrali bulalım} \\ \int 9\ x\ \mathrm{e}^{-3x}\ \mathrm{d}x &= -3x\ \mathrm{e}^{-3x} + \int 3\ \mathrm{e}^{-3x}\ \mathrm{d}x \\ &= -(3x+1)\ \mathrm{e}^{-3x} \end{aligned}$$

Sınırları koyarsak:

$$\begin{aligned} P(X < \frac{50}{60}) &= -(3x+1)\ \mathrm{e}^{-3x}\Big|_0^{5/6} \\ &= 0.7127 \qquad \blacksquare \end{aligned}$$

Not: Bunu 2 adet üstel dağılımlı RD'nin toplamını kullanarak da çözebilirdik. Nasıl peki?

X_1 ve X_2, $\lambda = 3$ hata/saat olan üstel dağılımlı R.D.'ler ise bize sorulan $P(X_1+X_2 < 5/6)$. Bu şekilde de devam edilebilir.

S2: **AYRIK** RASGELE DEĞİŞKENLER, X ve Y'nin "*Birleşik Olasılık KÜTLE Fonksiyonu*" aşağıdaki gibidir.

$$\mathrm{p}(x,y) = \begin{cases} \alpha\ x\ y \ , & 0 < x+y < 5, \qquad x,y \in \mathbb{Z} \\ 0 \ , & \text{diğer} \end{cases}$$

Aşağıdaki soruları cevaplayınız.

(a) `(5 p)` α'yı bulunuz.

Çözüm 2a:

Önce aralığa bakalım:

$$0 < x+y < 5 = 1 \le x+y \le 4$$

olur. Değer uzayının kısıtlamalarından biri de $\mathrm{p}(x,y) = \alpha\ xy \ge 0$ olması gerektiğidir, (α pozitif ise !!) bu durumda x ve y'nin değerleri 0'dan küçük olamaz (Not: α negatif de olabilir! Ama onu gözönüne almayalım). x'i serbest tutarsak y aşağıdaki şartı sağlar:

$$1 - x \le y \le 4 - x$$

$x \cdot y$'nin 0 olduğu yerleri dikkate almazsak y'nin alt limitini $1-x$'den değil 1'den başlatmamız mümkün olur. $x = 4$'te $4 - x = 0$ olacağı için x'in üst sınırı 3 alınabilir. OKF'nin aşağıdaki şartı da sağlaması gerekir:

$$
\begin{aligned}
1 &= \sum_{x\in R_x}\sum_{y\in R_y} \mathrm{p}(x,y) \\
1 &= \sum_{x=1}^{3}\sum_{y=1}^{4-x} \mathrm{p}(x,y) \\
1 &= \alpha\sum_{x=1}^{3}\sum_{y=1}^{4-x} x\ y
\end{aligned}
$$

Bunu tablo ile yapmakta fayda var. Değer uzaylarını gözönüne alarak aşağıdaki tabloya olasılıkları yazalım:

	x		
y	**1**	**2**	**3**
1	α	2α	3α
2	2α	4α	0
3	3α	0	0

Toplamın 1 olması lazım.

$$
\begin{aligned}
15\ \alpha &= 1 \\
\alpha &= \frac{1}{15} \qquad \blacksquare
\end{aligned}
$$

Not: α'nın işareti verilmediği için $\alpha < 0$ durumuna da bakmamız gerekirdi. Fakat bu sınavda gerekli değil.

(b) `(15 p)` X'in $Y = y$ şartı altında şartlı olasılık kütle fonksiyonunu bulunuz. (Tablo şeklinde olabilir)

Çözüm 2b:

$Y = y$ olma şartı altında X'in OKF'si aşağıdaki gibi bulunabilir:

$$P(X = x|Y = y) = \mathrm{p_X}(x|y) = \frac{\mathrm{p_{XY}}(x, y)}{\mathrm{p_Y}(y)}, \quad \mathrm{p_Y}(y) > 0 \textit{ olan yerlerde}$$

$\mathrm{p_Y}(y)$, y'nin marjinal dağılımıdır. Bu da :

$$\begin{aligned} \mathrm{p_Y}(y) &= \sum_{x \in R_x} \mathrm{p}(x, y) \\ &= \sum_{x=1}^{4-y} \alpha\ x\ y \qquad 1 \leq y \leq 3 \end{aligned}$$

Y'nin marjinal dağılımını bir önceki tablodan bulabiliriz. Aşağıya tekrar yazıp, sutünları ve satırları toplayarak marjinal dağılımları bulabiliriz.

	x			
y	**1**	**2**	**3**	$\mathrm{p}_Y(y)$
1	α	2α	3α	6α
2	2α	4α	0	6α
3	3α	0	0	3α
$\mathrm{p}_X(x)$	6α	6α	3α	

Yukardaki formülü şu şekilde uygulayabiliriz:

$\mathrm{p_X}(x|y)$	x			
y	**1**	**2**	**3**	$\mathrm{p}_Y(y)$
1	$\alpha/6\alpha$	$2\alpha/6\alpha$	$3\alpha/6\alpha$	6α
2	$2\alpha/6\alpha$	$4\alpha/6\alpha$	0	6α
3	$3\alpha/3\alpha$	0	0	3α

Burdan $P(X = x|Y = y) = \mathrm{p_X}(x|y)$ aşağıdaki tablodaki gibi çıkar:

$\mathrm{p_X}(x|y)$	x		
y	**1**	**2**	**3**
1	1/6	1/3	1/2
2	1/3	2/3	0
3	1	0	0

Bunu yazarken $P(X = x|Y = y) = \mathrm{p}_{\mathbf{X}}(x|y)$'nin yazılışına dikkat edin. $P(X = x|Y = y)$ şeklinde olması lazım, şartlı olasılık olduğunu gösteren $|$ sembolu çok önemli. $P(X = x, Y = y)$ ise şartlı OKF değildir, birleşik OKF'dir!!

(c) `(10 p)` X ve Y'nin kovaryansını ve korelasyon katsayısını bulunuz.

Çözüm (2c):

Kovaryans ve korelasyon katsayısı aşağıdaki gibi bulunur.

$$
\begin{aligned}
\mu_X = E(X) &= \sum_x x\, p_X(x) \\
&= 1 \cdot 6\alpha + 2 \cdot 6\alpha + 3 \cdot 3\alpha \\
&= 27\alpha \\
&= 27/15 = 7/5 = 1.8 \\
\mu_Y = E(Y) &= 27/15 = 1.8 \\
E(X^2) &= 1^2 \cdot 6\alpha + 2^2 \cdot 6\alpha + 3^2 \cdot 3\alpha \\
&= 57/15 = 19/5 = 3.8 \\
E(Y^2) &= 19/5 = 3.8 \\
V(X) = \sigma_X^2 &= E(X^2) - \mu_x^2 \\
&= 19/5 - 81/25 = 0.56 \\
V(Y) = \sigma_Y^2 &= 0.56 \\
E(XY) &= 1 \cdot 1 \cdot \alpha + 1 \cdot 2 \cdot 2\alpha + 3 \cdot 1 \cdot 3\alpha \\
&\quad + 2 \cdot 1 \cdot 2\alpha + 2 \cdot 2 \cdot 4\alpha \\
&\quad + 3 \cdot 1 \cdot 3\alpha \\
&= 43/15 = 2.86667 \\
\sigma_{XY} &= 2.86667 - 1.8 \times 1.8 \\
&= -0.37333 \qquad \blacksquare \qquad \text{Kovaryans} \\
\rho_{XY} &= \frac{\sigma XY}{\sigma_X\ \sigma_Y} \\
&= -0.6786 \qquad \blacksquare \qquad \text{Korelasyon (veya Korelasyon Katsayısı da deniyor)}
\end{aligned}
$$

(d) `(15 p)` $Z = g(X,Y) = X - Y$ ise, Z'nin olasılık kütle fonksiyonunu bulunuz. (Tablo şeklinde olabilir)

Çözüm 2d**:**

X ve Y nin değer uzayını kullanarak Z'nin değer uzayına bakalım.

Z	X		
Y	**1**	**2**	**3**
1	0	1	2
2	-1	0	
3	-2		

Burdan Z'nin OKF'si bulunr:

$$
\begin{aligned}
P(Z=-2) &= P(X=1, Y=3) = 3/15 \\
P(Z=-1) &= P(X=1, Y=2) = 2/15 \\
P(Z=0) &= P(X=1, Y=1) + P(X=2, Y=2) = 5/15 \\
P(Z=1) &= P(X=2, Y=1) = 2/15 \\
P(Z=2) &= P(X=3) = 3/15
\end{aligned}
$$

■

S3: Bir piyango çekilişinde ödül kazanma ihtimali %0.15 olsun. Ödül miktarı da 100 TL, bilet ücreti 1 TL olsun. Basılan bilet sınırsız sayılacak çok yüksek olduğunu varsayalım. Bir kişi bu biletlerden 150 bin adet satın alıyor.

(a) `(5 p)` Bu biletlerin ortalama kaç tanesi 100TL ödül kazanır?

Çözüm 3a**:**

X: 150000 biletten 100 TL ödül kazanan bilet sayısı ise, X Binomiyel R.D.'dir ve parametreleri $n = 150000$ ve $p = 0.0015$'dir.

$$P(X=x)=\binom{n}{x}\ p^x\ (1-p)^{n-x}$$

Bize sorulan ortalamadır. Binomiyel rasgtele değişkenin ortalaması np olduğuna göre:

$$E(X)=\mu=150000\times 0.0015=225 \qquad \blacksquare$$

Not: Binom dağılımında λ sembolünün kullanıldığı bir durum yoktur. Sadece Binom R.D.'sini Poisson R.D'sine benzetmede kullanılabiliyor fakat bu işlediğimiz şeyler arasında değil. Birçok öğrenci ortalama için μ sembolünü kullanmak yerine λ sembolünü kullanmış, 1 kişide olsaydı gözden kaçırmış denebilirdi ama çok kişide aynı hata olunca bu hatayı dikkate almak gerekiyor.

(b) `(10 p)` 100 TL Ödül kazanan bilet sayısının 200 ile 250 arasında olma ihtimali yaklaşık olarak nedir?

Çözüm 3b:

Bize sorulan $P(200<X<250)$'dir. Bunu da Binomiyel dağılım formülü ile yapmak istersek:

$$P(200<X<250)=\sum_{x=200}^{250}\binom{150000}{x}\ 0.0015^x\ (1-0.0015)^{150000-x}$$

Bunu çözmek zahmetli olacağı için bunu Standart Normal Dağılıma benzeterek yapabiliriz. Bu durumda, μ: X'in beklenen değeri, σ: X'in standart sapması olmak üzere,

$$Z=\frac{X-\mu}{\sigma}$$

yaklaşık olarak bir Standart Normal Dağılımlı R.D. olur. Ortalamayı önceki şıkda çözmüştük, onu ve standart sapmayı bulup yazalım.

$$\begin{aligned}\mu &= 225\\ \sigma &= \sqrt{np(1-p)}=\sqrt{224.6625}=14.9887\end{aligned}$$

Burdan:

$$
\begin{aligned}
P(200 < X < 250) &= P\left(\frac{200 - 224.6625}{14.9887} < Z < \frac{250 - 224.6625}{14.9887}\right) \\
&= P(-1.668 < Z < 1.668) \\
&= \phi(1.668) - \phi(-1.668) \\
&= 0.9525 - 0.0475 \\
&= 0.905 \qquad \blacksquare
\end{aligned}
$$

(c) `(10 p)` Bu biletlerin %1'ine de amorti çıktığını varsayalım, yani amorti çıkan biletlerin bilet ücretlerinin iade edildiğini varsayalım. Bu kişinin bu biletlere yaptığı harcamadan zarar etmeme ihtimali nedir?

Çözüm 3c:

Bu biletlerin $150000 \times 0.01 = 1500$ tanesine amorti çıkar, bunlar kendi parasını öder. Geriye $150000 - 1500 = 148\,500$ bilet kalır. Bunlara ödenen para 148500 TL'dir. Bu yüzden en az $148500/100 = 1485$ bilete ödül çıkması lazımdır ki zarar edilmesin.

O zaman, Y: $n = 148500$ bilet içinden ödül çıkan bilet sayısı, Binomiyel bir R.D., $p = 0.0015$ ise ve $\mu :$ Y'nin ortalaması, $\sigma :$ Y'nin standart sapması olmak üzere:

$$
\begin{aligned}
\mu = E(Y) &= n\ p = 148500 \times 0.0015 = 222.75 \\
\sigma = \sqrt{V(X)} &= \sqrt{n\ p\ (1-p)} = \sqrt{148500 \times 0.0015 \times 0.9975} = 14.9136 \\
P(Y > 1485) &= P(Z > \frac{1485 - 222.75}{14.916}) \\
&= P(Z > 84.624) \\
&= 1 - \phi(84.624)
\end{aligned}
$$

Standart Birikimli Normal Dağılım tablosundan $\phi(84.624)$'yi bulmaya çalıştığınızda tabloda bunun bulunmadığını göreceksiniz. Sağa doğru gittiğinizde birikimli dağılımın 1'ye yaklaştığını biliyoruz, bu yüzden bu değer de 1'e yaklaşır. Burdan:

$$
\begin{aligned}
P(Y > 1485) &= 1 - \phi(84.624) \\
&\approx 1 - 1 \\
&= 0 \qquad \blacksquare
\end{aligned}
$$

Bulunur.