پاسخ کوییز سوم

Dirichlet Process

$P \sim DP(\alpha) \longrightarrow \bar{\alpha}$ is base measure

سوال ۱-

(a) افراز $\{A, A^c\}$ را روی فضای S، measurable، در نظر بگیرید.

$$\longrightarrow P(A, A^c) \sim Dir(\alpha(A), \alpha(A^c)) = Beta(\alpha(A), \alpha(A^c))$$

می دانیم که امید ریاضی توزیع بتا برابر است با:

$$X \sim Beta(\alpha, \beta) \longrightarrow \mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$$

$$\Rightarrow \mathbb{E}[P(A)] = \frac{\alpha(A)}{\alpha(A)+\alpha(A^c)} = \frac{\lambda\bar{\alpha}(A)}{\lambda\underbrace{(\bar{\alpha}(A)+\bar{\alpha}(A^c))}_{1}} = \bar{\alpha}(A)$$

(b)

$$X \sim Beta(\alpha, \beta) \longrightarrow Var[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$\Rightarrow Var[P(A)] = \frac{\alpha(A)\,\alpha(A^c)}{(\alpha(A)+\alpha(A^c))^2(\alpha(A)+\alpha(A^c)+1)} = \frac{\lambda^2\,\bar{\alpha}(A)\bar{\alpha}(A^c)}{\lambda^2(\lambda+1)} = \frac{\bar{\alpha}(A)\bar{\alpha}(A^c)}{1+|\alpha|}$$

(c)

$$(P(A\cap B), P(A-B), P(B-A), P((A\cup B)^c)) \sim Dir(\alpha(A\cap B), \alpha(A-B), \alpha(B-A), \alpha((A\cup B)^c))$$

$$cov[P(A), P(B)] = cov[P(A\cap B)+P(A-B),\ P(A\cap B)+P(B-A)]$$

$$= Var[P(A\cap B)] + \underbrace{cov[P(A-B), P(A\cap B)]}_{} + \underbrace{cov[P(B-A), P(A\cap B)] + cov[P(B-A), P(A-B)]}_{}$$

cov of dirichlet

$$= \frac{\bar{\alpha}(A\cap B)(1-\bar{\alpha}(A\cap B))}{1+|\alpha|} + \frac{-\bar{\alpha}(A-B)\bar{\alpha}(A\cap B)}{1+|\alpha|} + \frac{-\bar{\alpha}(B-A)\bar{\alpha}(A\cap B)}{1+|\alpha|} + \frac{-\bar{\alpha}(B-A)\bar{\alpha}(A-B)}{1+|\alpha|}$$

$$(\bar{\alpha}(A\cap B) = \bar{\alpha}(A) - \bar{\alpha}(A-B) = \bar{\alpha}(B) - \bar{\alpha}(B-A))$$

$$= \frac{\bar{\alpha}(A\cap B) - \bar{\alpha}(A)\,\bar{\alpha}(B)}{1+|\alpha|}$$

سوال ۲ -
$$\mathrm{corr}(P(A_1), P(A_2)) = \frac{\bar{\alpha}(A_1 \cap A_2) - \bar{\alpha}(A_1)\bar{\alpha}(A_2)}{\sqrt{Var(P(A_1))\,Var(P(A_2))}} \overset{disjoint}{=} \frac{-\bar{\alpha}(A_1)\bar{\alpha}(A_2)}{\sqrt{\quad}} < 0$$

در فرآیند دیریشله، همبستگی هر دو مجموعه مجزا مقداری منفی است. یک رفتار در توزیع احتمالاتی تصادفی که مورد انتظار ما است این است که جرمی که به مجموعه های نزدیک یکدیگر نسبت داده می شود با هم افزایش و کاهش پیدا کنند. اما همبستگی منفی عکس این انتظار را نشان می دهد. در واقع می توان گفت در فرآیند دیریشله توپولوژی فضا برای نسبت دادن جرم ها درنظر گرفته نشده است.

سوال ۳ - برای یک جایگشت از ورود مشتری ها می توان نوشت

$$P(x_1, \ldots, x_n) = \frac{\left[\prod_{i=1}^{k} (n_i - 1)!\right] \alpha^{k-1}}{(\alpha+1)(\alpha+2)\cdots(\alpha+n-1)}$$

واضح است که این رابطه احتمال برای تمام جایگشت ها یکسان خواهد بود. پس می توان گفت هر دنباله n تایی exchangeable است. ⇐ infinitely exchangeable است

- هر دنباله یک دنباله iid exchangeable است. اما عکس این قضیه لزوماً درست نیست برای مثال رستوران چینی با ۲ مشتری را درنظر بگیرید.