باسمه تعالی

نظریه پیشرفته آمار

آزمون پایان‌ترم

زمستان ۱۴۰۱

---

۱) برآوردگر بیش‌ترین درست‌نمایی (MLE) چیست و به چه معنا (به صورت مجانبی) بهینه است؟ (اثبات لازم نیست.)

۲) نرخ کشف خطا (FDR) در آزمون فرض هم‌زمان چه مفهومی دارد؟ روش بنجامینی و هوشبرگ چیست و با چه فرضیاتی در این روش نرخ کشف خطا کنترل می‌شود؟ (اثبات لازم نیست.)

۳) یک آماردان نسبتاً برجسته روشی ابداع کرده که (با تعدادی فرض احتمالاتی/زمین‌شناسانه) به کمک داده‌های ماه‌های گذشته تخمین نااریبی از تعداد زلزله‌های ماه آینده به دست می‌دهد. متاسفانه مقدار تخمین‌گر گاهی منفی می‌شود! برای رفع این مشکل تخمین‌گر جدیدی می‌سازیم که هرگاه تخمین‌گر اولیه منفی بود آن را به صفر تغییر می‌دهیم. اریبی و واریانس دو تخمین‌گر را با هم مقایسه کنید.

۴) فرض کنید $\hat{y}_1, \cdots, \hat{y}_B$ تخمین‌گرهایی از y باشند. اگر $\hat{y}_i$ها واریانس σ^2 و هر دو تا از آن‌ها همبستگی ρ داشته باشند، واریانس تخمین‌گر $\hat{y} = \frac{1}{B}(\hat{y}_1 + \cdots + \hat{y}_B)$ با افزایش B چطور تغییر می‌کند؟ به کمک این نتیجه بررسی کنید که:

(a) زیاد کردن تعداد درختان در الگوریتم جنگل تصادفی چه اثری بر نتیجه دارد؟

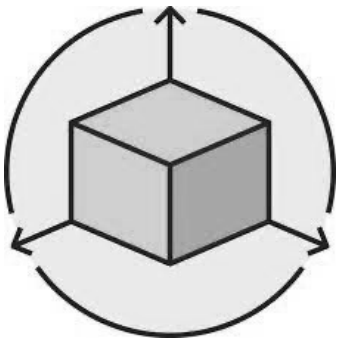
(b) کم یا زیاد کردن تعداد متغیرها که در هر مرحله از ساخت درختان جنگل به تصادف انتخاب می‌شوند و انشعاب روی یکی از آن‌ها انجام می‌گیرد چطور؟

۵) یک متغیر پاسخ y و تعدادی متغیر مستقل $x_1, \cdots, x_p$ داریم و فرض می‌کنیم وابستگی آن‌ها به شکل زیر است:

$$y = f_1(x_1) + \cdots + f_p(x_p) + \epsilon$$

که ϵ مستقل از x_iها با میانگین صفر است. اگر به اندازه‌ی کافی داده داشته باشیم در هر کدام از حالت‌های زیر روشی برای تخمین توابع f_i پیشنهاد کنید:

(a) تنها یک متغیر مستقل داشته باشیم ($p = 1$) و f_1 تابعی هموار (smooth) باشد.
(b) f_iها توابع خطی و تنها تعداد کمی از آن‌ها ناصفر باشند.
(c) f_iها توابع هموار باشند.
*(d) f_iها توابع هموار و تنها تعداد کمی از آن‌ها ناصفر باشند.



۶) در یکی از روش‌های اعمال دراپ‌اوت (dropout)، در فرآیند یادگیری شبکه‌ی عصبی، هر بار وزن‌ها را با اختلال‌های گوسی مستقل نویزی می‌کنیم! برای فهم بهتر این کار، یک مساله‌ی رگرسیون خطی با تابع زیان تصادفی زیر در نظر بگیرید:

$$L(\beta) = \frac{1}{n}\sum_{i=1}^{n}[y_i - \sum_{j=1}^{p}(\beta_j + \epsilon_j)x_{i,j}]^2$$

که ϵ_iها متغیرهای گوسی مستقل با میانگین صفر و واریانس λ هستند. امید گرادیان تابع زیان، $\mathbb{E}[\partial L(\beta)/\partial\beta]$، را محاسبه و تا جای ممکن ساده کنید. اگر از روش کاهش گرادیان تصادفی برای حل این مساله‌ی بهینه‌سازی استفاده کنیم به چه جوابی خواهیم رسید؟ این نتیجه معادل اضافه کردن چه هموارسازی (Regularization) به تابع خطای کمترین مربعات عادی است؟