

۱ الگوریتم PC

الگوریتم را مطابق شبه کد داده شده در متن تمرین پیاده سازی کردم.
برای تست استقلال از آزمون فرض روی نرخ همبستگی (در حالت شرطی همبستگی جزئی) استفاده کردم.

۲ قواعد MEEK

ابتدا با استفاده از اطلاعات خروجی تابع PC، مجموعه v-structure ها را مطابق زیر می یابیم. در واقع از مرحله S2 استفاده میکنیم، زیرا مرحله اول را در همان تابع PC انجام داده ایم.

- S1 Form an undirected graph G by the following rule.
 A is adjacent to B in G if and only if there does not exist a set $S \subseteq V \setminus \{A, B\}$ such that $\mathcal{M} \models A \perp\!\!\!\perp B | S$. If there is such an S let $Sep(A, B) = S$.
- S2 For all unshielded triples $\langle A, B, C \rangle$ orient $A \rightarrow B$ and $C \rightarrow B$ if $B \notin Sep(A, C)$.

قدم اول قواعد MEEK

سپس توسط قواعد زیر ساختار تلاش میکنیم CPDAG نزدیک به DAG واقعی را بیابیم.



- S1 Let Π_I be the result of phase I. Orient every edge which can be oriented by successive applications of rules R1, R2 and R3; i.e. close Π_I under rules R1, R2 and R3.

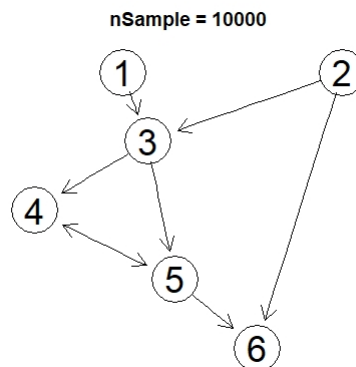
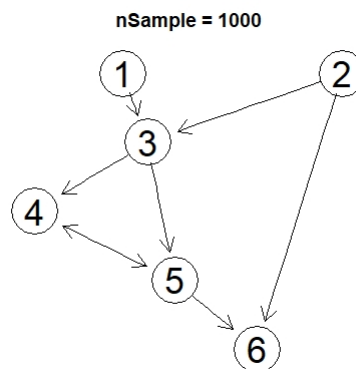
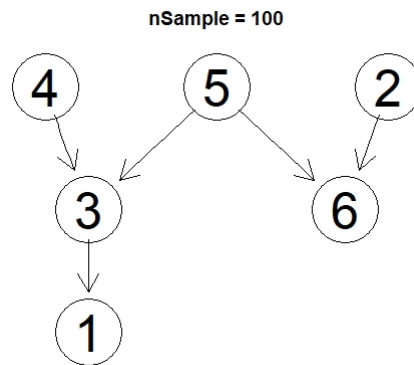
قدم دوم قواعد MEEK

تابع meek به منظور پیاده سازی روش MEEK نوشته شده است. (از آنجا که در مسئله ما اطلاعاتی پیشین از گراف نداریم اعمال R4 لزومی ندارد.)

۳ بررسی عملکرد الگوریتم

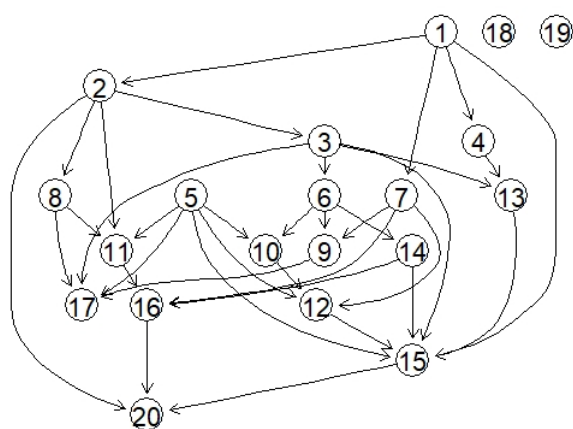
(الف)

داده را تولید میکنیم، سطح اطمینان را نیز برابر ۵ درصد میگیریم. برای تعداد نمونه‌های مختلف خروجی به صورت زیر است: همانگونه که مشاهده میشود در حالت ۱۰۰ نمونه جواب خوبی بدست نمی‌آید که این به دلیل عملکرد ضعیف آزمون فرض استقلال است که هم یالها و هم مجموعه v-structure ها درست تشخیص داده نمیشوند، به تبع آن قوانین MEEK درست عمل نمیکنند. در حالت تعداد نمونه ۱۰۰۰ و ۱۰۰۰۰ جواب بسیار خوب و نزدیک به DAG واقعی است و فقط یال ۴ به ۵ تشخیص داده نمیشود که البته قواعد MEEK این حالت را نمیتوانند تشخیص دهند. می‌توان ادعا کرد به CPDAG ماکسیمال رسیده‌ایم.



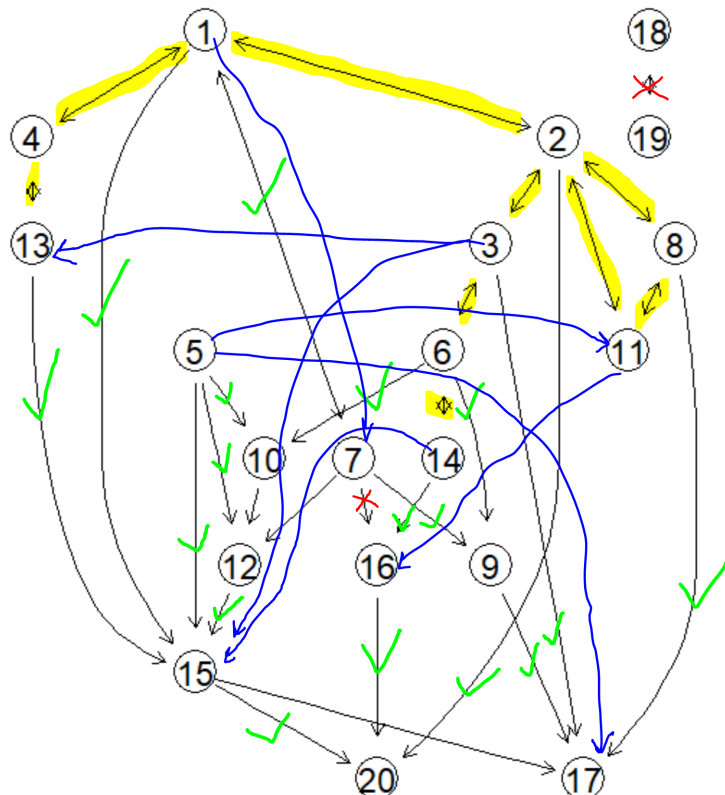
(ب)

در این قسمت همانگونه که گفته شد، گراف تصادفی میسازیم و عملکرد الگوریتم خود را بررسی میکنیم. ابتدا یک مثال را بررسی میکنیم و سپس نمودار خواسته شده را رسم میکنیم. گراف زیر را ساخته ایم.



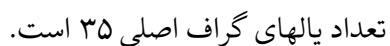
خروجی randomDAG

حال ابتدا با تولید ۱۰۰۰ نمونه از هر راس خروجی را مشاهده میکنیم. در این حالت فاصله همینگ خروجی از گراف واقعی برابر ۱۸ است. (گراف اصلی ۳۵ یال دارد) علت خطاها بیشتر به دلیل خطا در وجود یا عدم وجود یالهاست که نتیجه ناصحیح بودن آزمون فرض در الگوریتم PC است.

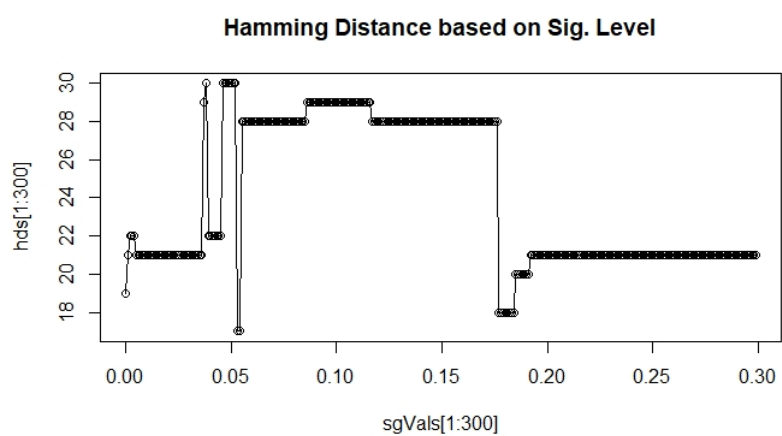
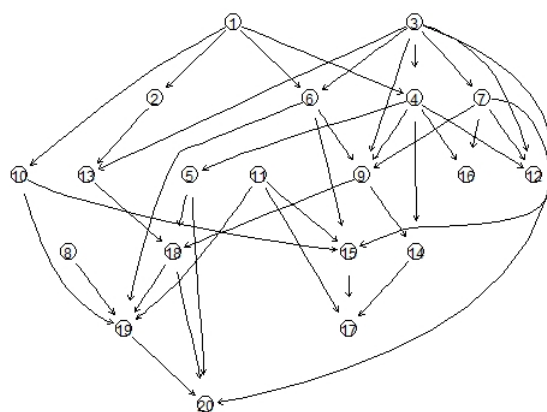


خروجی الگوریتم برای $n = 1000$ ، خطوط آبی: یال های تشخیص داده نشده، تیک سبز: جهت درست، فلش قرمز: جهت غلط، ضربدر قرمز: یال اضافی، زرد: جهت نامعلوم

برای همین گراف نمودار فاصله همینگ بر حسب بازه اطمینان به صورت زیر است.



५



تعداد یالهای گراف اصلی ۴۰ است.