1. **Na czym polega metoda superpozycji podczas generowania liczb pseudolosowych?**

Metoda superpozycji

Stosujemy tam, gdzie trudno odwrócić dystrybuantę, ale można ją podzielić na części, które łatwo odwracać.

$$F_X(x) = \sum_k c_k F_X^{(k)}(x) \qquad 0 < c_k < 1 \qquad \sum_k c_k = 1$$

Przykład:

$$f_X(x) = \frac{5}{12}(1 + (x-1)^4) \qquad x \in (0,2)$$

2. **Na czym polega metoda eliminacji podczas generowania liczb pseudolosowych?**

Metoda eliminacji: stosujemy gdy f(x) niezerowe tylko w przedziale (a,b) i ograniczone (przez stała <c). Metoda postępowania: losujemy punkt (x,y) w prostokącie o bokach (a,b) i (0,c), jeżeli leży on nad f(X) to akceptujemy x, w przeciwnym wypadku losujemy kolejny punkt.

3.Na czym polega metoda „z przekształceniem" podczas generowania liczb pseudolosowych?

Metody z przekształceniem – szukamy innej zmiennej losowej o tej samej dystrybuancie (funkcji gęstości prawdopodobieństwa).

Przykład: $f_X(x) = nx^{n-1} \qquad x \in (0,1)$

generujemy jako $x = max(\gamma_1, \gamma_2, \ldots, \gamma_n) \qquad \gamma_i \in (0,1)$

4. **Co nazywamy statystyką zmiennej losowej?**

Statystyka – funkcja zmiennych losowych obserwowanych w próbie, sama też jest zmienną losową.

5. **Co to jest estymator?**

Estymatorem parametru Θ nazywamy statystykę o rozkładzie prawdopodobieństwa zależnym od Θ i oznaczamy $T_n(\Theta)$ lub $T_n(x_1,x_2,\ldots,x_n;\ \Theta)$.

Zatem dla danego parametru może istnieć wiele estymatorów o różnych własnościach. Pożądanymi własnościami są:

1. Zgodność – estymator $T_n(\Theta)$ jest zgodny gdy

$$\forall \varepsilon > 0: \quad \lim_{n\to\infty} P(|\underbrace{T_n(\Theta)}_{\text{wartość estymatora (statystyki)}} - \Theta| < \varepsilon) = 1$$

2. Obciążenie (jego brak) – estymator $T_n(\Theta)$ jest zmienną losową (różne próby dają różne wartości estymatora), zatem ma wartość oczekiwaną i wariancję. Obciążenie estymatora to

$$B_n \equiv E(T_n(\Theta)) - \Theta$$

6. **Co to jest estymator zgodny?**

1. Zgodność – estymator $T_n(\Theta)$ jest zgodny gdy

$$\forall \varepsilon > 0: \quad \lim_{n\to\infty} P(|\underbrace{T_n(\Theta)}_{\text{wartość estymatora (statystyki)}} - \Theta| < \varepsilon) = 1$$

7. **Co to jest estymator nieobciążony?**

Estymator jest nieobciążony, gdy niezależnie od wielkości próby

$$\boxed{E(T_n(\Theta)) = \Theta \quad \text{czyli} \quad B_n = 0}$$

8. **Co to jest estymator asymptotycznie nieobciążony?**

Estymator jest asymptotycznie nieobciążony gdy

$$\lim_{n\to\infty} B_n = 0$$

9. **Co to jest estymator najbardziej efektywny?**

3. Efektywność – estymator jest **najbardziej efektywny** gdy ma najmniejszą wariancję.

Optymalnie estymator powinien być: zgodny, nieobciążony i najbardziej efektywny.

10. **Który estymator wartości oczekiwanej ma optymalne własności?**
11. **Co jest wynikiem estymacji punktowej?**

Estymacja punktowa to metoda statystyczna mająca na celu oszacowanie wartości badanego zjawiska. Chodzi tu o znalezienie konkretnej liczby (nie przedziału), która jest estymatorem parametru w populacji (w odróżnieniu od estymacji przedziałowej)

12. **Dlaczego, tam gdzie to możliwe, powtarzamy pomiar i liczymy średni wynik, a niezadowalamy się jednym pomiarem?**

Wiecej wyników = wieksza dokładność

13. **Podaj nieobciążony estymator wariancji.**

Nieobciążony estymator wariancji

$$S_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$$

14. **Co jest wynikiem estymacji przedziałowej?**

Estymacja przedziałowa – podanie przedziału liczbowego,

$$\left[T_n^L(\theta), T_n^P(\theta)\right]$$

wewnątrz którego, z założonym prawdopodobieństwem γ =1-α, leży prawdziwa wartość parametru Θ.

15. **Co to jest przedział ufności dla parametru ϴ?**

wewnątrz którego, z założonym prawdopodobieństwem g =1-a, leży prawdziwa wartość parametru ϴ.

$$P\left(T_n^L(\theta) \le \theta \le T_n^P(\theta)\right) = 1 - \alpha \equiv \gamma$$

16. **Od czego zależą wartości krańców przedziału ufności?**

2). Końce przedziału zależą od próby i od γ, a nie zależą funkcyjnie od Θ.

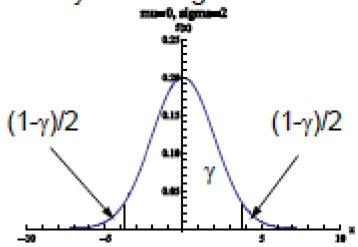
17. **Opisz procedurę estymacji przedziałowej.**

Aby znaleźć przedział ufności szukamy statystyki o znanym rozkładzie prawdopodobieństwa zależnej od estymowanego parametru Θ i próby.

Przypadek 1 – znamy σ(X)

Twierdzenie: Statystyka

$$Z \equiv \frac{\overline{X} - E(X)}{\sigma(\overline{X})} = \frac{\left(\overline{X} - E(X)\right)\sqrt{n}}{\sigma(X)}$$



ma rozkład normalny N(0,1). (Średnia ma rozkład N(E(X),σ^2(X)/n), standaryzacja przekształca rozkład normalny w normalny).

Chcemy znaleźć przedział ufności taki, że

$$P\left(Z^L \le Z \le Z^P\right) = \gamma$$

Kwantyle rozkładu N(0,1)

Jest to spełnione przez $Z^L = Z_{\frac{1-\gamma}{2}}$ $\quad i \quad$ $Z^P = Z_{\frac{1-\gamma}{2}+\gamma=\frac{1+\gamma}{2}}$

18. Wymień 3 metody szukania estymatorów.

-Metoda momentów
- Metoda największej wiarygodności
-Metoda najmniejszych kwadratów

19. Opisz procedurę szukania estymatorów w metodzie momentów.

Metoda momentów polega na porównaniu momentów względem początku układu współrzędnych $\widetilde{\mu}_k \equiv E(X^k)$ wyliczonymi wprost z definicji a wartościami ich estymatorów wyliczonymi na podstawie n-elementowej próby

$$\boxed{T_n(\widetilde{\mu}_k) = \frac{1}{n}\sum_{i=1}^{n}(x_i)^k}$$

czyli $\widetilde{\mu}_k(\theta_1, \theta_2, \ldots, \theta_p) = \frac{1}{n}\sum_{i=1}^{n}(x_i)^k$

Rozwiązanie tak otrzymanego układu p (k=1,....,p) równań na $\Theta_1, \Theta_2, \ldots, \Theta_p$ przyjmujemy za estymatory parametrów

Estymatory uzyskane metodą momentów są zgodne i asymptotycznie nieobciążone (czasami nawet nieobciążone).

20. Opisz procedurę szukania estymatorów w metodzie największej wiarygodności.

Metoda największej wiarygodności polega na szukaniu wartości parametrów $\Theta_1, \Theta_2, \ldots, \Theta_p$, dla których funkcja największej wiarygodności osiąga maksimum.
Tak otrzymane wartości przyjmujemy za estymatory parametrów $\Theta_1, \Theta_2, \ldots, \Theta_p$.

21. Zdefiniuj funkcję największej wiarygodności.

Funkcja największej wiarygodności

$$\widetilde{L}(\theta_1, \theta_2, \ldots, \theta_p) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_p)$$

22. **Opisz procedurę szukania estymatorów w metodzie najmniejszych kwadratów.**

Metoda najmniejszych kwadratów polega na szukaniu wartości parametrów $\Theta_1, \Theta_2, \ldots, \Theta_p$, dla których funkcja

$$\sum_{i=1}^{n} w_i \left(g(y_i; \theta_1, \theta_2, \ldots, \theta_p)\right)^2$$

osiąga minimum. Współczynniki liczbowe w_i określają wagę jaką przykładamy do kolejnych wartości y_i, mogą być to na przykład odwrotności kwadratów błędów pomiaru y_i.

Tak otrzymane wartości przyjmujemy za estymatory parametrów $\Theta_1, \Theta_2, \ldots, \Theta_p$.

23. **Co to jest funkcja regresji I rodzaju zmiennej Y względem zmiennej X ?**

Funkcją regresji I rodzaju zmiennej Y względem zmiennej X nazywamy warunkową wartość oczekiwaną E(Y|X) traktowaną jako funkcję zmiennej X.

$$\begin{pmatrix} E(Y \mid X = x) = \int_{-\infty}^{+\infty} y f_{Y|X}(y \mid x) dy \\ \\ f_{Y|X}(y \mid x) = f_{X,Y}(x,y) / f_X(x) \end{pmatrix}$$

24. **Co to jest liniowa funkcja regresji II rodzaju ?**

Liniowa funkcja regresji II rodzaju przybliża liniowo E(Y|X)

$$E(Y \mid X) \approx u(x) = ax + b$$

25. **Co to jest test statystyczny ?**

Testowanie hipotez statystycznych pozwala na sprawdzenie na podstawie wyników próby, przy zadanym poziomie ufności, czy jakieś twierdzenie (hipotezę) dotyczące populacji generalnej jest prawdziwe. Taką procedurę nazywamy **testem statystycznym**.

26. **Co to jest hipoteza zerowa ?**

Jest to hipoteza poddana procedurze weryfikacyjnej, w której zakładamy, że różnica między analizowanymi parametrami lub rozkładami wynosi zero. Przykładowo wnioskując o parametrach hipotezę zerową zapiszemy jako:

$$H_0 \colon \theta_1 = \theta_2$$

27. **Co to jest hipoteza prosta**

jednoznacznie określają rozkład, funkcję gęstości prawdopodobieństwa lub dystrybuantę zmiennej losowej

28. Co to jest błąd I rodzaju i jak wiąże się z poziomem istotności testu ?

29. Co to jest błąd II rodzaju ?

Określenie obszaru krytycznego na podstawie a jest niejednoznaczne. Można to poprawić wprowadzając prawdopodobieństwo **błędu II-go rodzaju**: przyjęciu fałszywej H0, gdy w rzeczywistości prawdziwa jest hipoteza alternatywna H1. Prawdopodobieństwo to oznaczamy b, a 1-b nazywamy **mocą testu**.

30. Opisz procedurę testu statystycznego np. do zbadania H_0: E(X)=X_0.

Zał. Próbka o liczebności n pochodzi z rozkładu N(E(X),σ^2).

Badamy H_0: $E(X)=X_0$.

Statystyką testową jest:

Gdy znamy odchylenie standardowe σ: $U = \frac{(\overline{X} - E(X))\sqrt{n}}{\sigma(X)}$

Ma ona rozkład N(0,1).

Gdy nie znamy odchylenia standardowego σ: $t = \frac{(\overline{X} - E(X))\sqrt{n}}{S(X)}$

Ma ona rozkład t-Studenta o n-1 stopniach swobody.

H_1	W (znane σ)	W (nieznane σ)
$E(X) \neq X_0$	$\lvert u \rvert > u_{1-\alpha/2}$	$\lvert t \rvert > t_{1-\alpha/2}$
$E(X) > X_0$	$u > u_{1-\alpha}$	$t > t_{1-\alpha}$
$E(X) < X_0$	$u < u_{\alpha}$	$t < t_{\alpha}$